

P R O J E C T M E M O R A N D U M

RAND

*A Classification System for
Inpatient Rehabilitation
Patients: A Review and
Proposed Revisions to the
Functional Independence
Measure-Function Related
Groups*

*Grace M. Carter, Daniel A. Relles,
Joan L. Buchanan, with Donna Bean,
Tenzing Donyo, Kenneth Rosenfeld,
Lisa Rubenstein, Karen L. Spritzer*

PM-682-HCFA

September 1997

Prepared for the Health Care Financing Administration

RAND/UCLA/Harvard Center for Health Care Financing Policy Research

This document is unreviewed and unedited.

PREFACE

This is a report on the case classification system that underlies the construction of a model Prospective Payment System for Medicare inpatient rehabilitation cases. The model payment system is described in:

Grace M. Carter, Joan L. Buchanan, Tenzing Donyo, Moira Inkelas, and Karen L. Spritzer, *A Prospective Payment System for Inpatient Rehabilitation*, RAND, PM-683-HCFA, 1997.

This report and the companion report cited above constitute the final report for delivery order number 9 under Master Contract Number 500-92-0023.

The classification system proposed here is based heavily on the Functional Independence Measure--Function Related Group (FIM-FRG) classification system developed by Margaret Stineman and colleagues at the University of Pennsylvania and SUNY-Buffalo. This group continues work on FIM-FRGs, and Dr. Stineman was kind enough to specify the clinical categories used in the latest version. We have used these categories in this report.

We used data obtained from the Adult FIM Database based on the Uniform Data System for medical rehabilitation (UDSmrSM), which was created and is maintained by the Uniform Data System for Medical Rehabilitation (UDSMR), a division of University of Buffalo Foundation Activities, Inc. This division operates within the Center for Functional Assessment Research (CFAR), under the Department of Rehabilitation Medicine of the School of Medicine and Biomedical Sciences of the State University of New York at Buffalo. The UDSmrSM contains discharge records from participating hospitals, and these discharge records include the Functional Independence Measure (FIMSM), a service mark of the UDSMR.

UDSmr owns all copyrights to the FIMTM measurement instrument (i.e., the Uniform Data Set, the Functional Independence Measure scale, the related impairment codes and the Guide for Use of the Uniform Dataset for Medical Rehabilitation), and to all databases of patient

data compiled using the FIMTM measurement instrument. FIMTM is UDSmr's trademark used in connection with its measurement instrument and databases; FIMSM is UDSmr's service mark used in connection with its data management service. Appropriate licenses are available from UDSmr to qualified health facilities, researchers and consultants.

All copyrights and related intellectual property rights in the Function Related Group ("FRG") classification system, including trade secrets and proprietary information ("FRG Information"), are owned by The University of Pennsylvania. UDSmr has the exclusive worldwide license to use the FRGs and the FRG Information for the development, creation, sale, distribution, and use of products or services incorporating any portion of the foregoing.

CONTENTS

Preface.....	iii
Figures.....	vii
Tables.....	ix
Summary.....	xi
Acknowledgments.....	xv
1. INTRODUCTION	1
2. SUMMARY OF PANEL FINDINGS	4
Clinical and Functional Status Panel.....	4
Patient Classification System Design	5
Adequacy of the Functional Independence Measure	6
Possible Future Improvements to the FIM-FRGs	8
Classification and Payment System Design Panel.....	9
Study Population and Data Quality	9
Classification Analysis	10
3. DATA	12
UDSmr.....	12
Medicare Data.....	14
Matching Discharges from MEDPAR and UDSmr.....	15
Matching Algorithm	15
Quality of the Match	20
Representativeness of UDSmr.....	22
Hospital Characteristics	22
Patient and Stay Characteristics	24
4. METHODS	29
Overview.....	29
Variable Definitions.....	31
Dependent Variables	31
Independent Variables	32
Sample Sizes.....	33
Statistical Models.....	35
5. RESULTS	39
How Many FRGs?.....	39
Modeling Readmissions, Interrupted Stays, and Comorbidities....	43
The Chosen Model and Its Structure.....	45
Evaluation of Selected Model Structure.....	53
Model Accuracy	53
Explanatory Power	57
FRG Selection.....	60
6. CONCLUSIONS	67

Appendix

A. REPLICATION OF 1990 FRGs	69
B. COMORBIDITIES	72
C. GOLD STANDARD MODEL COEFFICIENTS AND T STATISTICS	78
D. MEMBERS OF TECHNICAL ADVISORY PANELS	85
References	87

FIGURES

3.1.	Distribution of All Hospitals by Percentage of Their MEDPAR Records Matched to a UDSmr SM Record	21
3.2.	Distribution of Hospitals with a Full Year of Data by Percentage of Their MEDPAR Records Matched to a UDSmr Record	21
3.3.	Percentage of MEDPAR Rehabilitation Cases in Sample Credentialed UDSmr Facilities in 1994, by State	25
5.1.	Average LOS in UDSmr FRGs, All UDSmr Cases	50
5.2.	Average LOS in Medicare LOS FRGs, Medicare Cases	50
5.3.	Average Cost in Medicare Cost FRGs, Medicare Cases	51
5.4.	Average Cost for All UDSmr Cases in UDSmr FRGs, for Fitting and Evaluation Samples	61
5.5.	Average LOS for Medicare Cases in UDSmr FRGs, for Fitting and Evaluation Samples	61
5.6.	Average Cost of Medicare Cases in UDSmr FRGs, for Fitting and Evaluation Samples	62
5.7.	Scatter Plot of UDSmr FRG Values of Mean LOS in Evaluation Sample Against Mean LOS in Fitting Sample, All Cases	62
5.8.	Scatter Plot of UDSmr FRG Values of Mean LOS in Evaluation Sample Against Mean LOS in Fitting Sample, Medicare Cases ...	63
5.9.	Scatter Plot of UDSmr FRG Values of Mean Cost in Evaluation Sample Against Mean Cost in Fitting Sample, Medicare Cases ..	63
A.1.	Comparison of Average LOS in 1994 with Average LOS in 1990, by 1990 FRG	70

TABLES

3.1.	Marginal Contribution of Dichotomous Variables to Match Score.....	18
3.2.	Marginal Contribution of Continuous Variables to Match Score.....	19
3.3.	Percentage of UDSmr SM Discharges that Match MEDPAR Discharges, by Medicare Payer Group.....	22
3.4.	Comparison of Number of UDSmr Hospitals with Number of Medicare Rehabilitation Hospitals, by Type of Hospital.....	23
3.5.	Mean Value for Various Characteristics of All Medicare and UDSmr Hospitals.....	24
3.6.	Patient Characteristics for Medicare Rehabilitation Inpatients, by UDSmr Status and Disabled Status.....	26
3.7.	Comparison of Resource Use for Medicare Rehabilitation Inpatients, by UDSmr Status and Disabled Status.....	27
4.1.	Independent Variable Definition.....	33
4.2.	Sample Size by RIC.....	34
4.3.	Description of Models Used in Analysis, by Research Question.....	38
5.1.	Sample Size and Number of Groups in Best CART Fit.....	40
5.2.	R-Square Statistics from CART Models, by RIC and Number of Groups.....	41
5.3.	Mean LOS by FRG.....	47
5.4.	Coefficients on Interrupted Stays in Regression Models Controlling for FRG and Comorbidities.....	51
5.5.	Coefficients on Comorbidities in Regression Models Controlling for FRG and Interrupted Stay.....	52
5.6.	Percentage of Errors in Minimum-Node CART Model.....	54
5.7.	Description of Model Structure and R-Squares.....	58
5.8.	Effects of Using All-Payer FRGs on Medcost and MedLOS Datasets.....	65
A.1.	Replication Results: R-Squared.....	71
B.1.	Comorbidity Variables.....	76
C.1.	Regression Coefficients from Gold Standard Models.....	79
C.2.	T Statistics from Gold Standard Models.....	82

SUMMARY

We used a combination of qualitative and quantitative methods to assess the suitability of the Functional Independence Measure--Function Related Groups (FIM-FRGs) as a basis for a Prospective Payment System (PPS) for Medicare inpatient rehabilitation. The current method of paying for this care is based on the Tax Equity and Fiscal Responsibility Act (TEFRA) of 1982 and provides fewer cost control incentives than a PPS would. Inpatient rehabilitation was exempted from the acute-care PPS because its patient classification system did not explain resources well for rehabilitation cases.

The FIM-FRGs were developed by Stineman et al. (1994) using 1990 data for a sample of rehabilitation hospitals and were shown to predict variation in length of stay (LOS) adequately. The FIM-FRGs are organized around clinically defined Rehabilitation Impairment Categories (RICs) such as stroke and traumatic brain dysfunction. The RICs were subdivided into FRGs based on motor and cognitive subscales of the Functional Independence Measure and age.

METHOD AND DATA

Expert Panels

We solicited expert opinion about the ability of the FIM to measure functional status and the suitability of the FRGs to be the unit of payment in a rehabilitation PPS. We held meetings with two different panels of experts, reviewed the literature with them, and received their opinions in discussions and in written survey responses.

Data

Two data systems are the primary sources for our analyses. The Uniform Data System for medical rehabilitation (UDSMr) records the FIM at admission and demographic information about rehabilitation patients in participating hospitals. We used data from approximately 140,000 adult cases discharged during calendar year 1994 from 352 hospitals. We also used Medicare discharge files (MEDPAR) that record demographic,

clinical, and charge information and hospital-level files providing facility characteristics and financial information.

We linked records from the MEDPAR and the UDSmr that described the same discharge. We were able to find a MEDPAR record for 94 percent of UDSmr cases where Medicare was listed as the primary payer. We matched 80 percent or more of MEDPAR cases to a UDSmr case for 78 percent of hospitals that participated in UDSmr throughout 1994.

Quantitative Analyses

Our quantitative analyses (1) examined the stability of the original FRGs by using them to predict 1994 data; (2) extended the FRGs to take account of previously unanalyzed types of cases (readmissions and interrupted stays) and newly available data (comorbidities and complications from Medicare data); and (3) evaluated the performance of FRGs when data are based on cost rather than length of stay and when data are restricted to Medicare cases.

In expanding and refitting the FRGs, we maintained the original structure in which RICs are partitioned to maximize explained resource use. We evaluated alternative models based on their statistical power and the amount of bias in predictions for cases grouped by demographic and hospital characteristics. This allowed us to examine the explicit trade-offs between statistical accuracy and other attributes such as simplicity and use of non-Medicare data.

Three methodologies were used in developing the alternative models: (1) Classification and Regression Trees (CART), (2) Ordinary Least Squares regression (OLS), and (3) a combination of CART and OLS. The dependent variables in the model were cost and LOS for Medicare patients and LOS for all payer adult UDSmr cases. The independent variables included FIM motor score, FIM cognitive score, age, and indicators for readmission, interrupted stays, comorbidities, and poverty.

FINDINGS

Panel Recommendations

Our advisory panels generally endorsed development of a rehabilitation PPS as a method of improving the poor incentives that exist in the current payment system. But they cautioned that it will be

important to verify the reliability of FIM scoring. They also suggested that the cognitive component of the FIM should be improved and possibly expanded and that better cognitive items could become important additional predictors of resource use. In addition, they recommended clarification of rules for assignment of impairment codes.

Representativeness of UDSmr Data

Although UDSmr contains only 24 percent of the hospitals that serve Medicare inpatient rehabilitation patients, those hospitals serve 40 percent of all Medicare rehabilitation hospital cases. The UDSmr overrepresents freestanding rehabilitation hospitals and underrepresents exempt units. The patient characteristics of the UDSmr sample of Medicare cases appear very similar to those of all Medicare rehabilitation patients. Disabled recipients in the sample also are similar to all Medicare disabled rehabilitation recipients.

FIM-FRG Analysis

FIM-FRGs defined by CART analysis are robust, effective predictors of resource use. FRGs based on motor FIM score, cognitive FIM score, and age were stable over time. Although there was a decline in average LOS between 1990 and 1994, the 1990 FRGs defined by Stineman et al. explain 1994 LOS just as well as they explain LOS in their original 1990 data. FRGs fit on all payer UDSmr cases were as effective in predicting both the LOS and cost of Medicare cases as were groups defined based on Medicare data alone. To reduce the number of cases where small changes in score would cause changes in reimbursement, we recommend that the number of FRGs per RIC be limited to at most five. When this limit is imposed, the explanatory power of the FRGs is roughly equal to that of the fifth revision of the Diagnosis Related Groups (DRGs) (34 percent) and exceeds the explanatory power of DRGs for medical cases (18 percent).

We find that cases with interrupted stays and, in many RICs, comorbidities or complications cost substantially more than other cases in the same FRG. Although CART analysis defines FRGs that predict resource use, it does not adequately capture these main effects. Thus, we suggest using a RIC-specific multiplier for the weight of such cases.

Readmissions were found to differ little from other cases in the same FRG.

We could not reliably measure the effect of poverty at the case level. However, a rehabilitation PPS could provide a facility-level adjustment for facilities to account for the distribution of such patients across facilities.

Our definition of the comorbidity variable was derived empirically. It would be worthwhile for the Health Care Financing Administration (HCFA) to examine whether it is appropriate to pay for all the conditions identified.

Despite these limitations, we believe that the revised FIM-FRG described here is an adequate patient classification system for use in a rehabilitation PPS.

ACKNOWLEDGMENTS

We thank our project officer, William Buczko, Ph.D., of the Health Care Financing Administration's Office of Research and Demonstrations, for his support throughout the project. He helped us secure the data necessary for the research, participated in our advisory panels, provided a detailed critique of early versions of our final report, and helped fill us in on the policy background for our study.

We also thank each of the members of our advisory panels. Their names and affiliations are listed in Appendix D. They engaged in long, intense, sometimes conflicted sessions in which we learned much about measuring functional status and about the requirements of a payment system for inpatient rehabilitation.

This work would not have been possible without the foundation built by Dr. Carl Granger and his colleagues at SUNY-Buffalo and Dr. Margaret Stineman and her colleagues at the University of Pennsylvania. Their initial work on the FIM, the FIM-FRGs, and the UDSmr is the foundation of our research.

We also thank all those who reviewed our early drafts. Our RAND colleague, Donna Farley, made numerous useful suggestions for clarifications and additions. We received thoughtful detailed reviews from several panel members and from Drs. Granger and Stineman. Of course, any mistakes are our responsibility.

1. INTRODUCTION

Inpatient rehabilitation, exempted from Medicare's Prospective Payment System (PPS), is an important and growing component of care for Medicare beneficiaries. In FY 1993, Medicare paid 3.7 billion dollars for such care--a 95 percent increase from 1990 (ProPAC, 1996). This growth was on top of a 250 percent increase in the period from 1985 to 1990 (ProPAC, 1993). Medicare payment policy during this time frame attempted to constrain the growth in inpatient rehabilitation expenditures, but inadvertently created payment inequities between older and newer facilities that may have spurred the formation of new providers. Thus, both policymakers and the older, established inpatient rehabilitation provider community have acknowledged a need for payment policy revisions that could include a payment system similar to PPS but based on a new patient classification system for inpatient rehabilitation.

Payment for inpatient rehabilitation care of Medicare beneficiaries is currently made under the Tax Equity and Fiscal Responsibility Act (TEFRA) of 1982. The payment amount depends on a per-case target amount that is calculated from historical costs at the facility trended forward and on the hospital's actual cost per case. If costs are less than the target amount, then the hospital receives its costs plus a incentive payment. If costs are greater than the target amount, then the hospital receives its costs minus a penalty. Because the target amounts have not kept up with inflation, older facilities are at a disadvantage relative to newer facilities that have target amounts closer to current costs. In addition, because the system contains no case-mix adjustment, it penalizes hospitals whose case mix increased more than the average. The lack of a case-mix adjustment also creates incentives to limit access and/or care for patients requiring intensive rehabilitation.¹

The primary reason for the initial exclusion of rehabilitation hospitals from the PPS was that Diagnosis-Related Groups (DRGs) could

¹Further details about TEFRA and discussion of its problems may be found in the introduction to the companion report, Carter et al. (1997).

not predict resource use at these facilities very well. Functional status, measured by activities of daily living, communication skills, and cognition, is more correlated with rehabilitation charges for the patient than diagnoses alone (Hosek et al., 1986).

The development of a patient classification system for case-mix adjustment was hindered by the lack of an appropriate and widely accepted functional status measure for inpatient rehabilitation. The Functional Independence Measure (FIMSM) was developed to fill this need (Hamilton et al., 1987). It was subsequently adopted by many inpatient rehabilitation providers. Fiedler et al. (1996) report that 410 hospitals from 48 states reported FIM data in 1994 to the Uniform Data System for medical rehabilitation (UDSMrSM). The FIM is an 18-item measure covering six domains: self-care (six activities of daily living), sphincter control (two items on bowel and bladder management), mobility (three transfer items), locomotion (two items on walking/wheelchair use and stairs), communication (two items on comprehension and expression), and social cognition (three items on social interaction, problem solving, and memory). The first four domains--self-care, sphincter control, mobility, and locomotion--are often combined into a single motor scale. Similarly, the last two domains--communication and social cognition--are often combined to form a single cognitive scale. All 18 items are scored into one of seven levels of function ranging from complete dependence (1) to complete independence (7).

Using the FIM, a team of researchers from the University of Pennsylvania has developed a patient classification system, Function Related Groups, referred to as the FIM-FRGs (Stineman et al., 1994). The FIM-FRGs are organized around 18 (now 20) clinically defined Rehabilitation Impairment Categories (RICs) such as stroke, traumatic brain dysfunction, and nontraumatic spinal cord injury.² The original 18 RICs were subdivided into 53 patient classes using the FIM motor score, FIM cognitive score, and age. These classes were defined using a

²The full set of RICs will be defined in Section 4 (Table 4.2).

recursive partitioning algorithm that grouped patients using lengths of stay.

The Health Care Financing Administration (HCFA) contracted with RAND for an independent review and evaluation of FIM-FRGs using more recent data and the initial design of a possible PPS for inpatient rehabilitation based on FIM-FRGs. The review and evaluation of FIM-FRGs are described in this report, along with the proposed revisions to the FIM-FRGs that arose from this activity. The design of a model rehabilitation PPS is reported in a companion paper, Carter et al. (1997).

The original development of the FIM-FRGs was a research effort and therefore Stineman et al. did not deal with several issues required for implementation. For example, they deleted a large number of unusual cases in order to create a homogeneous dataset for use in developing the FIM-FRGs.

We here evaluate the FIM-FRGs for their suitability as the basis of a payment system. Our quantitative analyses (1) examined the stability of the original FRGs by using them to predict 1994 data; (2) extended the FRGs to take account of previously unexamined cases (readmissions), previously unused information (interrupted stays), and newly available data (comorbidities and complications from Medicare data); and (3) evaluated the performance of FRGs when cost rather than length of stay (LOS) is used to form groups and when only Medicare cases rather than all cases are used to form groups.

The advice of two technical advisory panels was used in the research. The first panel reviewed the adequacy of the FIM and the second panel provided guidance on the development of the patient classification system. Panel recommendations and findings are summarized in Section 2, followed by a description of the data used in the project (Section 3) and the analytic methods (Section 4). The review findings and recommendations are presented in the last two sections of the report.

2. SUMMARY OF PANEL FINDINGS

Two expert panels helped us with the qualitative evaluation of the suitability of the FIM-FRGs for use in a rehabilitation PPS (RPPS). The Clinical and Functional Status Panel reviewed the development of the FIM and the patient classification system, the FIM-FRGs. Panel topics included (1) how currently available data can best be used in the development of a patient classification system for reimbursement, (2) how to monitor such a payment system and create proper incentives, and (3) the adequacy of the FIM and the FIM-FRGs and how they might be improved. Although the panelists were asked to comment on many aspects of a rehabilitation payment system design, their primary focus was on clinical and functional assessment issues.

The second panel was the Classification and Payment System Design Panel. Its primary focus was on extensions and/or redesign of the patient classification system and on issues in the design of the RPPS. It reviewed the development of the FIM-FRGs, our initial replication of this effort using more recent data, and payment issues.

To assist each panel, RAND prepared a technical notebook with background materials from the literature. Where feasible, we supplemented the notebook with preliminary analyses of the project data. The panel findings and recommendations are a synthesis of both the meeting discussions and responses to post-meeting questionnaires. The post-meeting questionnaires included a series of statements, and panelists were asked to rate their level of agreement on a 5-point scale ranging from strongly agree to strongly disagree.

CLINICAL AND FUNCTIONAL STATUS PANEL

Our first panel was composed of experts in functional assessment; it included two physicians from the rehabilitation community and two social scientists. The membership of the panel is listed in Appendix D. All of the panelists had conducted research in the functional assessment area and had developed assessment measures. Three of the four panelists were affiliated with alternative functional assessment measurement

systems. As a consequence of this selection process, our panelists were well informed on issues in functional assessment, but often had strong vested interests that made it difficult to achieve consensus. Thus, although the panel was effective in pointing out areas of the FIM and FIM-FRGs that require further research, they may not have always rendered objective judgments on the importance of these issues.

Patient Classification System Design

The Clinical and Functional Status Panel agreed that impairment groups together with a set of functional status measures to modify them are clearly the right organizing concepts for a rehabilitation patient classification system. The group tended to agree that the major impairment groups currently identified in the UDSmr were the correct ones. They also agreed that the impairment groups recently added to the UDSmr such as debility and developmental disability should be analyzed for inclusion as rehabilitation impairment categories in a FIM-FRG system. The group also discussed whether the World Health Organization's International Classification of Impairments, Disabilities, and Handicaps (ICIDH) system was conceptually a more appealing organizing foundation for developing impairment groups than the International Classification of Diseases, 9th Revision, clinical modification (ICD-9-CM) system. It was agreed that operationalizing and implementing an ICIDH-like coding system was not practical at this stage.

The panel generally agreed that, with the current set of RICs, some cases will provide opportunities for alternative assignment, such as a patient with neck pain resulting from osteoarthritis who could be classified in an arthritis category or in the pain syndrome category, thus encouraging "gaming" (upcoding) if these categories had different payment weights. Physicians saw this as a more important issue than nonphysician panelists.

Physicians on the panel also strongly agreed that further splitting of the FIM/FRG RICs based on additional etiologic diagnoses might become important as RICs if evaluated using a larger sample size (e.g., there is only one cardiac category).

The ability of rehabilitation facilities to correctly code cases using ICD-9 codes was of concern to some panelists. We were told that diagnostic coding on the UDSmr had improved over time. There was some speculation that units attached to PPS hospitals might do a better and more sophisticated job of coding than freestanding facilities.¹ Concern was also expressed that classification and coding would be cleaner for the 10 Medicare-approved rehabilitation diagnoses than for other impairment groups.

Adequacy of the Functional Independence Measure

The Clinical and Functional Status Panel found that the concepts addressed in the FIM motor scale are clinically meaningful aspects of rehabilitation care and the actual 13 motor-scale items capture these concepts well. Panelists agreed, but less strongly, that the 13 motor-scale items cover the most important motor-domain determinants of resource use. The panelists had inconsistent views about whether there were important ceiling and floor effects for motor items in the inpatient rehabilitation populations. Ceiling or floor effects would be important if there was substantial variation in performance that was related to resource use for patients who received the highest (ceiling) score or lowest (floor) score. In the discussion, there was general agreement that the FIM motor items were limited in their usefulness to inpatient settings because ceiling effects precluded use in outpatient settings.

The panelists were less satisfied with the FIM cognitive scale than with the motor scale. Although they believed these items also address clinically meaningful aspects of rehabilitation care, they thought that individual items do not measure the concepts well. They did not agree that the five cognitive items covered the most important cognitive domain determinants of resource use. Panelists agreed that there are significant ceiling effects in the cognitive scale items, but had inconsistent views on whether there were also important floor effects in these items.

¹Our subsequent analysis suggests that this speculation was correct. See Carter et al. (1997).

Missing Items. FIM scoring rules do not allow for missing or unscored items. If clinicians are unable to score an item, they are instructed to score the patient in the lowest performance category (unable to perform). Several of the panelists felt strongly that this was not an appropriate way to handle missing data. In particular, it is likely to bias admission scores downward as missing values are perceived to be more common at admission.

Reliability. Existing studies of the reliability of the FIM were all done by UDSmr personnel. Three of the four panelists strongly agreed that additional reliability studies by independent evaluators were needed before any decisions were made about adopting the FIM for payment purposes. Panelists held differing views about whether a standardized protocol for data collection should be adopted before FIM reliability could be adequately established.

Standardized Assessment Procedures. Panelists agreed that functional status assessment for payment purposes should be based on clinical observation of performance (as FIM instructions state) rather than on the rater's assessment of the patient's capacity to perform. However, they held differing views on the adequacy of the standardization of FIM assessment procedures. Concern over the lack of standardization had two components. First, who completes the assessment is not standardized, so in some institutions individual items are scored by different clinicians with differing types of training, or the scoring may be done by a multidisciplinary rehabilitation assessment team. In other institutions, a single individual with unspecified training may complete the form. The implications of variability in assessment administration are unknown and probably warrant additional study. Further, one panelist noted that it may not make sense to ask physical and occupational therapists who consistently assess on a capacity basis to score the FIM on a performance basis, since these individuals do not regularly observe routine performance. The second area of concern was that scoring some of the items, particularly in the cognitive area, gave the rater a great deal of discretion in terms of what evidence was used in the assessment and/or how often a behavior had occurred.

Training. Panelists were clear that training procedures for functional assessment must not be based only on written scenarios. Further, periodic retesting and retraining are integral to maintaining data quality. The social scientists on the panel indicated that the training procedures for the FIM were not adequate, well supported, or standardized. Despite this observation, panelists generally agreed that the average rehabilitation hospital could accurately score the FIM. They held differing views on whether adequate screens for data quality at the patient level existed and whether there were adequate means of ensuring institutional competence in functional assessment.

Possible Future Improvements to the FIM-FRGs

FIM-FRGs currently use only three classification variables to subdivide the RICs: the admission FIM motor-scale score, the admission FIM cognitive-scale score, and patient age. Panelists indicated that FIM-FRGs would benefit from the inclusion of additional clinical indicators of illness/impairment severity. Their suggestions included a chronic disease list, body system count, the Charleson comorbidity index, and an approach similar to what Freeman et al. (1995) proposed for DRGs which would classify all comorbidities into discrete groups (e.g., unimportant, minor, major). Other variables mentioned by panelists were gender, social support, and time from onset to initiation of rehabilitation. All agreed that these were somewhat problematic, but some panelists felt that they should perhaps be explored to determine their importance.

We asked panelists to suppose that we could freely augment or change the FIM/UDSmr data system and to tell us what they would change. One of the panelists felt strongly that the system should be implemented as is. The other three panelists indicated that better measures for the following concepts could be developed: (1) problem solving, (2) attention/concentration, (3) oral comprehension, (4) endurance/deconditioning, and (5) social support. Additional concepts mentioned by at least two panelists include (1) producing written language, (2) orientation and short-term memory, (3) long-term memory, (4) producing spoken language, (5) social interaction, (6) ability to get into and out

of an automobile, (7) living situation, (8) knowledge of medications, (9) degree of motor loss, (10) spasticity/involuntary movement, (11) joint limitations, and (12) sensory deficit.

Clinical concepts that panelists felt needed to be added to the functional assessment include measures of (1) depression, (2) decubiti/skin condition, and (3) recent acute myocardial infarction. As above, one of the panelists felt strongly that nothing should be added or changed, so this recommendation is based on a consensus of the remaining three. Additional concepts mentioned by at least two panelists include (1) nutritional status, (2) dialysis, and (3) home oxygen.

CLASSIFICATION AND PAYMENT SYSTEM DESIGN PANEL

Our second panel consisted of five members--experts in rehabilitation medicine, patient classification, and payment system design. The members are listed in Appendix D.

Study Population and Data Quality

As will be discussed in more detail in Section 3 of this report, the UDSmr contains abstracts of rehabilitation discharges from participating facilities. The Classification and Payment System Panel agreed that the UDSmr data appeared to be of acceptable quality for the construction of a classification system that would form the basis of payment for inpatient rehabilitation.

To help panelists understand the effects of using all adult discharges in the UDSmr population or restricting it to only Medicare beneficiaries, we provided a preliminary analysis that showed the effects of performing the Classification and Regression Trees (CART) analysis on the whole sample and on the Medicare sample only.² For a subset of the RICs, particularly those relating to traumatic injuries, the percentage of variance explained was noticeably lower when the data were restricted to the Medicare sample. With the exception of one member, the panel agreed that the population base for the development of

²A final version of this analysis may be found in Section 5.

the classification system should include non-Medicare patients to help ensure the system's broadest applicability.

With the exception of one member, the panel agreed that the development of the classification system should be limited to 1994 data because of the larger sample of hospitals and the likelihood of improved and consistent coding.

Classification Analysis

In preparation for the Classification and Payment System Design Panel meeting, we performed a preliminary CART analysis to replicate and extend the original developmental work. We looked at (1) how well the original FRGs performed on 1994 data, (2) how a new CART analysis performed when restricted to the original number of classifications, (3) how much improvement was obtained when all RICs could have up to five splits, and (4) how much improvement was obtained when all RICs could have up to ten splits.³

When we redid the CART analysis on the 1994 data and restricted the number of splits within each RIC to the number obtained in the original analysis, the order that variables entered into the tree was usually similar. Cut points, however, did differ across the years. With the exception of one panelist, there was agreement that the original set of FIM-FRGs showed good stability when replicated on the 1994 data.

Panelists also indicated that further analyses need not retain the technical details implemented in the original CART analysis. In fact, several panelists felt that we should consider different methodological approaches, including continuous regression models rather than patient classes. The panel was not able to agree on the best analytic approach. Panelists did agree that sensible stopping rules that recognize the larger volume of data available for our analyses should be implemented.

In judging the models in the preliminary analyses, panelists generally preferred the CART model that allowed five splits in each RIC to the model that limited the number of splits to the original level. In particular, additional splits were valuable in RICs 10-15, which were

³The first two analyses are described in Appendix A. Final versions of the last two analyses may be found in Section 5.

relatively smaller with poorer fits in the original data. Panelists were in agreement that the CART model that allowed ten splits per RIC did not perform significantly better than the model limited to five splits.

Four of five panelists agreed that each FRG should be separated from adjacent FRGs by an adequate range on the FIM separating variables to limit opportunities to game or leapfrog categories. However, there was no agreement on what constituted an adequate range. Further, there was no agreement on the importance of limiting the overall number of categories.

Four of five panelists also agreed that the original definition of outliers for exclusion from the classification analysis (cases with logarithm of length of stay (LOS) three standard deviations above the mean for the RIC) should be retained.

Four of five panelists also agreed that interrupted stays, readmissions, and transfers should be entered as potential classification variables in the CART analysis. Again, with one exception, panelists agreed that short stays for evaluation should be treated as a separate group or groups for classification and payment.

The advisory panel believed that, after accounting for case mix as measured by FIM-FRGs, comorbidities, and social support, remaining differences between freestanding rehabilitation hospitals and exempt units would arise from practice pattern differences and should not be compensated by HCFA.

Panelists indicated that FIM data should be used in whatever format resulted in the best explanatory power. The structure of groups and their splits need not be consistent across RICs. They also agreed that Rasch weights should be used if they contributed to the explanatory power of the models.⁴

With the exception of one panelist, there was agreement that performance-based measures such as the change in FIM score should be evaluated further for inclusion.

⁴Rasch transformation was not used in our analyses for the reasons cited in Section 4.

3. DATA

Two data systems--the Uniform Data System for medical rehabilitation (UDSmr) and Medicare program information--are the primary sources for our analyses. The UDSmr records functional status and demographic information for rehabilitation discharges from participating hospitals. The Medicare data files include discharge files recording demographic, clinical, and financial information and hospital-level files providing facility characteristics and financial information.

We describe our two data sources further below. We linked records from the Medicare Provider Analysis and Review (MEDPAR) and the UDSmr that described the same discharge (we describe the linking process and the quality of the resulting match in the third subsection below). The last subsection describes the extent to which the UDSmr participants are representative of the universe of Medicare patients.

UDSmr

The UDSmr describes rehabilitation stays in participating hospitals. The data include demographic descriptions of the patient (birth date, gender, zip code, ethnicity, marital status, living setting), clinical descriptions of the patient (condition requiring rehabilitation, ICD-9-CM diagnoses, FIM at admission, FIM at discharge) and the hospitalization (hospital, admission date, discharge date, charges, payment source, and an indicator of whether this is the first rehabilitation hospitalization for this condition, a readmission, or a short stay for evaluation). Although we did not use these data, the UDSmr system also contains follow-up information providing, for some cases, the FIM and living setting following discharge. UDSmr is operated within the Center for Functional Assessment Research, under the Department of Rehabilitation Medicine of the School of Medicine and Biomedical Sciences of the State University of New York at Buffalo as a division of the University at Buffalo Foundation Activities, Inc.

We used data from calendar year 1994. UDSmr provided us with data on 139,360 cases from 352 inpatient rehabilitation hospital facilities.

(Because some UDSmr facilities did not wish to participate, these numbers are lower than those reported by UDSmr.) Some of the cases are "interrupted stays," which occur when the patient's rehabilitation treatment is interrupted by an acute-care hospitalization and the person returns to the rehabilitation facility within 30 days after leaving the rehabilitation facility. Typically, Medicare data treat an interrupted stay with a single acute-care hospitalization as two rehabilitation discharges, one with two acute-care hospitalizations as three rehabilitation discharges, etc. Using these Medicare rules, the 139,360 cases correspond to 144,719 discharges.

The UDSmr contains the dates of transfer to and return from acute care, for up to three acute-care hospitalizations. As discussed later, we used this information to expand the UDSmr case records into their corresponding discharges in order to match with the Medicare data. We then aggregated the discharges back up to cases (after summing costs from each corresponding discharge) for the classification analyses reported here.

UDSmr provides training to participant hospitals and examinations that credential hospitals. Until a facility passes the credentialing exam, the data are not considered acceptable and were excluded from our data file. Because hospitals joined the UDSmr throughout 1994, we have only partial-year data from 59 of the 352 hospitals.¹ The data are routinely screened for out-of-range values and returned for correction. Our frequencies confirmed the accuracy and completeness of these screens. The success of our match to Medicare discharge records testifies further to the accuracy of the demographic information on the UDSmr. Unfortunately, diagnosis codes are not part of UDSmr's screening and frequently contain invalid values.

The only data problem we encountered concerned the occurrence of duplicate records describing the same case but usually with a different patient identifier. There were almost 2000 cases with duplicate records that were alike on almost all characteristics. We used only the last

¹For 4 of the 59 hospitals, the data cover only the early part of the year. We are not sure why these hospitals stopped sending data to UDSmr, but some may have failed their periodic recertification by UDSmr.

record received at UDSmr for the 85 percent of duplicate cases in which UDSmr was able to identify the date received. We used the last record in the original file sent by UDSmr for the remaining 15 percent of duplicate cases. Because there were so few records with doubt about the last record and because the duplicate records were so likely to be similar (for example, only 3.5 percent of the duplicate cases had different answers to any FIM question), we are convinced these records did not have a measurable impact on our analysis.

MEDICARE DATA

We used two Medicare data files. The MEDPAR file describes all inpatient stays (including rehabilitation stays) paid for by Medicare, and the Health Care Provider Cost Report Information System (HCRIS) provides facility information at the hospital and exempt unit level.

The MEDPAR file contains one record for each discharge (including from rehabilitation facilities) paid by Medicare. We used the 1994 calendar year MEDPAR, which matches the time period of our UDSmr data. The data include patient demographics (age, gender, race, residence zip code), clinical characteristics (diagnoses and procedures), and hospitalization characteristics (admission date, discharge date, days in intensive-care wards, charges by department, and reimbursement information).

Hospital cost reports are contained in the HCRIS. The cost report files contain information on facility characteristics, utilization data, and cost and charge data by cost center. Each record covers a hospital fiscal year, and all the hospital fiscal year records that began during a specific federal fiscal year are kept together in the same file. The file is named according to the number of years from the beginning of the acute-care PPS. We used the latest available files--the PPS 10 files (i.e., hospital fiscal years that began during the period from 10/1/92 through 9/30/93). Thus, the cost report data cover a somewhat earlier time period than the MEDPAR data, with the time difference varying by provider--from 15 months (for hospitals whose fiscal year begins on October 1) to, typically, six months (for hospitals whose fiscal year begins July 1).

The Exempt Hospital and Excluded Unit dataset contains HCRIS data for hospitals and units excluded from the PPS. The Minimum Dataset is an HCRIS file that contains data on all Medicare certified facilities, but only at the entire facility level. We took hospital location, number of beds,² and cost-to-charge ratios from the Exempt Hospital and Excluded Unit Dataset, supplemented by similar data from the PPS 10 Hospital Minimum Dataset for hospitals in Maryland and those paid under the acute-care PPS.

MATCHING DISCHARGES FROM MEDPAR AND UDSmr

Matching Algorithm³

We attempted to match each MEDPAR record from a hospital participating in the UDSmr to a corresponding UDSmr record. We expanded UDSmr records for interrupted stays so that there was one record for each discharge, with discharge dates corresponding to transfers to acute care (or final discharge) and admission dates corresponding to returns from acute care (or initial admission).

The UDSmr does not contain the Medicare beneficiary identifier, and therefore we needed to use a probabilistic matching algorithm based on characteristics of the beneficiary and the hospitalization. The first step in linkage was to determine the Medicare provider number corresponding to each UDSmr facility code.

The UDSmr provided us with a preliminary crosswalk of their facility codes with Medicare provider numbers. However, we needed the Medicare provider number used by the facility in 1994 (which is the provider number appearing on the MEDPAR), and there have been many changes in provider numbers and names over time because of reorganizations, purchases, and mergers. Provider numbers that did not appear in the Medicare cost report (HCRIS) files were candidates for review as were all those where no (or very few) matches were found. Our

²For a small number of exempt providers, number of beds was missing on the cost report file but was recovered from a printed extract of the provider of services file. When the available data indicated a provider had 0 beds, we assumed that data element was missing.

³The reader who is not interested in technical details will want to skip the next few pages to the subsection entitled "Quality of the Match."

first step was to search the cost report files for the hospital names reported by UDSmr. We (or UDSmr staff) contacted the hospital for remaining unresolved cases and asked for their 1994 Medicare provider number. This resolved almost all cases. However, there are two hospitals that we believe did not have the right provider number, because these providers had few MEDPAR rehabilitation records (0 and 2) despite having a large number of UDSmr records for which Medicare was listed as payer.

We insisted on a perfect match on hospital identity, and therefore searched for matches only within the same hospital. For freestanding hospitals (both rehabilitation and long-term care), we attempted to match all MEDPAR records to a UDSmr record. For acute-care hospitals, we attempted to match only records that either had a provider code of "T" which indicates that the case was cared for in a rehabilitation unit exempt from the PPS, or a DRG of 462, which indicates that rehabilitation was the primary reason for the hospitalization.

In addition to hospital identity, we used six variables to link the records: admission date, discharge date, zip code, age at admission, sex, and race.

The theory underlying the matching algorithm is given in Jaro (1995). The method considers each possible pairing of records (within a given rehabilitation facility) and computes the log of the odds of a match conditional on the values found on each file for each of the six linkage variables. It then selects the pairs with the greatest likelihood of being correct matches. A cutoff such that scores below the cutoff are not accepted as probable matches is chosen empirically.

The log of the odds of a match is computed as the sum of the marginal contributions of information for each of the six linkage variables (thus assuming independence across errors in the linkage variables). Three of the linkage variables (sex, race, and zip code) were treated dichotomously as either the same on both files or different.⁴ If one of these variables agreed on both files, the contribution of that variable to the log odds of a match for a

⁴Race was transformed into a binary variable indicating white or nonwhite.

particular combination of one record from each file is an estimate of the logarithm of

$$\frac{\text{Pr}(\text{the pair of records is a match} \mid \text{agreement})}{\text{Pr}(\text{pair does not match} \mid \text{agreement})}$$

Similarly, if a variable differed between the files, the contribution of that variable to the log odds is an estimate of the logarithm of

$$\frac{\text{Pr}(\text{the pair of records is a match} \mid \text{disagreement})}{\text{Pr}(\text{pair does not match} \mid \text{disagreement})}$$

In the case of a continuous variable, the marginal contribution to the linkage is based on the absolute value of the difference in the value of the variables on the two files:

$$\frac{\text{Pr}(\text{the pair of records is a match} \mid \text{difference})}{\text{Pr}(\text{pair does not match} \mid \text{difference})}$$

The first step in calculating these odds ratios was to take pairs thought to be sure matches and calculate the marginal error rates for each variable. This reveals the inconsistency of each variable between the two files. To determine these rates, for each linkage variable we successively matched all records in the two files on the remaining five variables and tabulated differences for the one held-out variable. For example, we would calculate the error rate in sex as the proportion of perfect matches on zip code, age, admission date, discharge date, and race, where the sex variable differed between the two files.

These error rates give us an estimate of the probability of finding disagreement given a correct match (and therefore also the complementary probability of finding agreement given a correct match). The probability of disagreement (and agreement) given no match were calculated based on randomly paired records from each file. We assumed that all these randomly paired records were nonmatches.

The last step was to compute the contribution to the log odds given agreement or difference. Assuming a prior probability of match that does not vary among candidate pairs, Bayes' theorem was applied to calculate the probabilities in the odds ratios that provide the marginal contribution of each variable. Dealing with continuous variables posed

additional problems because of the prior expectation that the distribution of log odds ratios would be somewhat smooth; in these cases, we smoothed the log odds using running averages of length 5. In the case of age, we noted additionally that there were spikes at 10, 20, and 30 (because errors in the first digit of year of birth are more likely than errors in both digits). We left those numbers intact rather than smoothing them with running averages.

Tables 3.1 and 3.2 show the contributions to cumulative scores posed by each of the six variables used in matching. The contributions are in base 10 logarithms and they only apply when both variables are nonmissing; otherwise, the contribution is zero. For continuous variables (age, admission date, discharge date), the difference is computed and bounded at 35; every difference larger than 35 receives the match score corresponding to a difference of 35.

After summing the marginal contribution of each of the six variables for each possible pairing, we sorted the pairings. We selected the highest scoring pair as a good match and then eliminated all other pairs containing either of the matched records in this pair. We continually repeated this process, selecting the highest scoring of the remaining feasible records and eliminating all pairs containing a matched record. We stopped including matches when the match score was below 2.00. This is a very conservative criterion, and indicates that any pairs that we accepted are likely to be correct, in the sense that agreement between the two records from each file is so strong that it is not likely to result from chance.

Table 3.1
Marginal Contribution of Dichotomous
Variables to Match Score

Agreement	Sex	Race	Zip Code
Same	0.27	0.08	1.24
Different	-1.47	-0.60	-0.79

Table 3.2
Marginal Contribution of Continuous Variables
to Match Score

Difference	Age	Admission	Discharge
		Date	Date
0	1.47	1.72	1.76
1	0.04	-0.50	-0.14
2	-0.92	-1.07	-0.73
3	-1.20	-1.00	-0.92
4	-1.39	-1.36	-1.10
5	-1.48	-1.25	-1.22
6	-1.60	-1.47	-1.49
7	-1.74	-1.53	-1.56
8	-1.81	-1.47	-1.64
9	-1.93	-1.39	-1.53
10	-1.44	-1.43	-1.47
11	-2.09	-1.49	-1.52
12	-2.15	-1.56	-1.57
13	-2.23	-1.67	-1.68
14	-2.25	-1.72	-1.73
15	-2.24	-1.75	-1.79
16	-2.22	-1.75	-1.75
17	-2.21	-1.74	-1.72
18	-2.20	-1.74	-1.73
19	-2.20	-1.83	-1.76
20	-1.64	-1.92	-1.76
21	-2.29	-2.03	-1.79
22	-2.34	-2.09	-1.84
23	-2.40	-2.13	-1.85
24	-2.43	-2.25	-1.82
25	-2.39	-2.25	-1.89
26	-2.36	-2.27	-1.92
27	-2.39	-2.26	-1.93
28	-2.46	-2.23	-1.98
29	-2.52	-2.08	-1.95
30	-1.47	-2.08	-1.91
31	-2.51	-2.15	-1.97
32	-2.53	-2.16	-2.03
33	-2.47	-2.38	-2.10
34	-2.44	-2.43	-2.16
35	-2.36	-2.52	-2.23

Quality of the Match

We matched 100,754 MEDPAR records to a corresponding UDSmr record. Several thousand of these discharges corresponded to part of an interrupted stay, so these matched to 97,679 original UDSmr records.

Figure 3.1 shows the distribution of the 352 UDSmr hospitals by the percentage of their MEDPAR records that matched a UDSmr record. (The hospital with no MEDPAR record is counted in the 0 cell.) Although 165 (47 percent) of the hospitals matched more than 90 percent of their records and 232 (66 percent) matched more than 80 percent of their records, there are a number of hospitals with a low match rate, primarily because their UDSmr data file covered only part of the year. Figure 3.2 restricts the sample to hospitals having a complete year of UDSmr data. Here the match looks better because almost all the hospitals with less than a 50 percent match rate have been excluded, as well as several in the 50 percent to 70 percent range. Among the hospitals with a full year of UDSmr data, 56 percent exceed a 90 percent match rate, 78 percent exceed 80 percent, and 98 percent exceed 50 percent. The single hospital in the 0 to 10 percent cell is the one where we believe that we do not know the correct provider number.

The match rate for cases at all hospitals with a full year of data is 85.6 percent overall. This match rate is noticeably lower for long-term-care hospitals (66.3 percent) than for any of the three other types of hospitals (freestanding rehabilitation, exempt unit, and other acute care are 86.2, 86.9, and 87.0, respectively). Our companion report, Carter et al. (1997), provides data on DRGs and charges for therapy that suggest that many of the cases at long-term-care hospitals that were not sent to UDSmr are not rehabilitation cases.

Another way to look at the quality of the match is to examine the proportion of UDSmr records that match to the MEDPAR. UDSmr, however, contains many cases not paid by Medicare. Table 3.3 shows the match rate by the UDSmr information on payer.

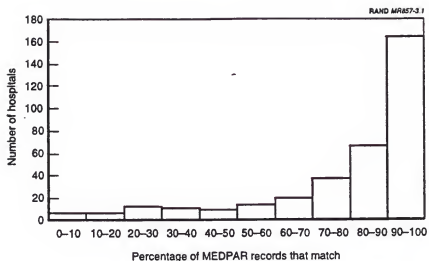


Fig. 3.1 Distribution of All Hospitals by Percentage of Their MEDPAR Records Matched to a UDSmrSM Record

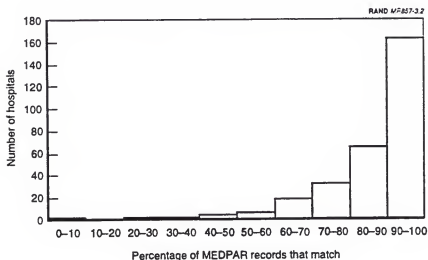


Fig. 3.2 Distribution of Hospitals with a Full Year of Data by Percentage of Their MEDPAR Records Matched to a UDSmr Record

Table 3.3
Percentage of UDSmrSM Discharges that Match MEDPAR Discharges
by Medicare Payer Group

Payer Group	Number of UDSmr Discharges	Percent Matched
Medicare primary payer	104,892	93.7
Medicare secondary payer	3,052	41.9
Medicare not payer	36,775	3.4

REPRESENTATIVENESS OF UDSmr

We next address the extent to which the hospitals participating in the UDSmr are representative of the set of all hospitals that provide inpatient rehabilitation care to Medicare beneficiaries and the extent to which UDSmr Medicare patients are representative of all Medicare rehabilitation patients.

Hospital Characteristics

UDSmr contains data from 352 inpatient hospital facilities. Table 3.4 divides these hospitals into four groups: (1) freestanding hospitals that received exemption from the PPS as rehabilitation facilities, identified using the third and fourth digit of their Medicare provider number (30); (2) exempt rehabilitation units of acute-care hospitals identified either by the existence of a rehabilitation sub-provider on the Exempt Hospital and Excluded Unit dataset or by a majority of their rehabilitation MEDPAR cases (cases assigned to DRG 462 or cared for in an exempt unit--i.e., with a provider type code of T) indicating care in an exempt unit; (3) other freestanding hospitals, primarily long-term-care hospitals (third and fourth digit of their Medicare provider number is 20); (4) other acute-care hospitals, including acute-care hospitals in Maryland as well as PPS hospitals.⁵

It is straightforward to define the universe of Medicare rehabilitation hospitals in groups 1 and 2. For groups 3 and 4, we wish to restrict the comparison to hospitals that provide care for patients with rehabilitation as their principal reason for hospitalization. We

⁵Maryland has a state waiver from the Medicare PPS. Thus, we cannot tell whether their acute-care hospitals would or would not meet the criteria for having an exempt rehabilitation unit.

Table 3.4

Comparison of Number of UDSmr Hospitals with Number of Medicare Rehabilitation Hospitals, by Type of Hospital

Type of Hospital	UDSmr	All	Percentage
		Medicare	UDSmr
Freestanding rehabilitation	98	209	47
Exempt unit of acute-care hospital	232	864	27
Other freestanding hospital	14	53	26
Other acute-care hospital	8	348	2
Total	352	1474	24

restrict the set of non-UDSmr long-term-care and "other acute-care" hospitals to those with cases in DRG 462. This probably accurately captures the universe of "other acute-care" hospitals. Although this captures about half of all long-term-care hospitals, it might underestimate the number treating cases for whom rehabilitation is the primary reason for hospitalization. Indeed, this criterion would miss several of the 14 long-term-care hospitals that participate in the UDSmr. (The UDSmr hospitals were all included as part of the universe of Medicare rehabilitation hospitals.)

As shown in Table 3.4, the UDSmr overrepresents freestanding rehabilitation hospitals and underrepresents exempt units. Although the UDSmr contains few hospitals in the "other acute-care" category, these care for a very small percentage of Medicare rehabilitation patients (2 percent).

Although UDSmr contains only 24 percent of the hospitals that serve Medicare inpatient rehabilitation patients, those hospitals serve 40 percent of all Medicare rehabilitation hospital cases. Table 3.5 shows that, on average, UDSmr participant hospitals care for more Medicare cases annually than non-UDSmr hospitals within each hospital group. The combination of the UDSmr tendency toward larger hospitals and the higher representation of freestanding rehabilitation hospitals--the category with the most patients per hospital--results in the UDSmr hospitals serving over twice as many Medicare rehabilitation cases as the average rehabilitation facility. Table 3.5 also shows that teaching hospitals are slightly overrepresented in the UDSmr.

Table 3.5
Mean Value for Various Characteristics of All Medicare
and UDSmr Hospitals

Hospital Characteristic	UDSmr	non-UDSmr	All Medicare
Medicare rehabilitation cases			
Freestanding rehabilitation	587.2	414.4	495.5
Exempt unit of acute-care hospital	278.8	219.8	235.7
Other freestanding hospital	258.6	40.7	98.3
Other acute-care hospital	231.4	13.1	18.1
All types	362.8	170.2	216.2
Rehabilitation beds ^a			
Freestanding rehabilitation	71.5	66.3	68.9
Exempt unit of acute-care hospital	28.6	24.4	25.6
Percent of hospitals with physical medicine residency			
Freestanding rehabilitation	11.2	5.4	8.1
Exempt unit of acute-care hospital	9.1	8.5	8.7
Other freestanding hospital	14.3	12.8	13.2
Other acute-care hospital	12.5	2.1	2.3
All types	9.9	6.4	7.3

^aNumber of rehabilitation beds is not available for the other hospital categories. Bed data are missing for 29 freestanding rehabilitation facilities (14 percent) and for 43 exempt units (5 percent).

Figure 3.3 shows that UDSmr hospitals account for a substantial portion of the Medicare rehabilitation cases in almost all states. Only Wyoming, Utah, New Mexico, and Maine are not represented. There are some UDSmr facilities in each region, although the southeast and mountain states appear to be slightly underrepresented.

Patient and Stay Characteristics

In order to examine the patient characteristics of Medicare rehabilitation patients in UDSmr hospitals, we used the matched MEDPAR-UDSmr discharge sample. Thus, this analysis reflects the effects of the partial-year sample available in some UDSmr hospitals as well as the sampling of hospitals.

RAND MR87-3.3

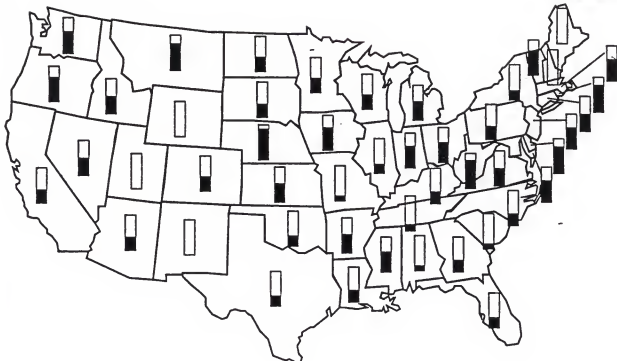


Figure 3.3-Percentage of MEDPAR Rehabilitation Cases in Sample Credentialed UDSmr Facilities in 1994, by State

Table 3.6 compares demographic characteristics of all Medicare rehabilitation patients with the matched UDSmr sample. The first three data columns cover all Medicare beneficiaries, while the last three cover only non-aged, disabled Medicare beneficiaries. In all the characteristics examined, the UDSmr sample of discharges appears very similar to the population and the UDSmr disabled sample of discharges is similar to all Medicare disabled discharges. Although differences are quite small, UDSmr patients are more likely to be white and less likely to die in the hospital than other rehabilitation patients.

Table 3.7 compares resources used per stay for UDSmr Medicare beneficiaries with those for other Medicare rehabilitation patients. Average length of stay for UDSmr cases is one day longer than for non-UDSmr patients (a 5.6 percent increase). This is explained in large part by the greater than expected proportion of UDSmr patients in freestanding rehabilitation hospitals and the longer expected LOS in

Table 3.6
Patient Characteristics for Medicare Rehabilitation Inpatients, by UDSmr Status and Disabled Status

Patient Characteristics	All Beneficiaries			Disabled		
	UDSmr	non-UDSmr	All Medicare	UDSmr	non-UDSmr	All Medicare
Average age	75.0	75.1	75.1	52.3	52.2	52.3
Male (%)	38.0	37.2	37.5	53.8	54.1	54.0
White (%)	88.2	85.8	86.6	74.6	69.9	71.3
Black (%)	8.3	10.6	9.9	19.9	23.9	22.6
Other race (%)	3.5	3.4	3.5	5.5	6.2	6.1
SSI recipient (%)	6.0	7.0	6.7	17.2	17.1	17.1
In-hospital death (%)	0.26	0.49	0.42	0.14	0.43	0.34
Sample size	100,445	216,511	316,956	7,662	17,300	24,962

SOURCE: Patient characteristics are all defined from the MEDPAR file. In-hospital death determined from the discharge destination field.

Table 3.7

**Comparison of Resource Use for Medicare Rehabilitation Inpatients, by UDSmr Status
and Disabled Status**

Patient Characteristics	All Beneficiaries			Disabled		
	UDSmr	non-UDSmr	All Medicare	UDSmr	non-UDSmr	All Medicare
Length of stay, days	18.8	17.8	18.1	20.1	18.8	19.2
Total charges, \$	19,169	19,019	19,066	21,701	21,549	21,596
Daily therapy charges, \$	319	309	312	314	288	296
Sample size	100,445	216,511	316,956	7,662	17,300	24,962

SOURCE: Patient characteristics are all defined from the MEDPAR file. In-hospital death determined from the discharge destination field.

such facilities. The average LOS of UDSmr patients in freestanding rehabilitation hospitals is 19.96 days compared to 19.83 days in freestanding non-UDSmr hospitals; the average LOS of UDSmr patients in exempt units is 17.33 days compared to 17.07 days in similar non-UDSmr hospitals. Despite the LOS differences, charges for UDSmr hospitals are similar to charges from non-UDSmr hospitals. Resources used by disabled beneficiaries are represented as well by the UDSmr as are resources used by all beneficiaries.

4. METHODS

OVERVIEW

Based on the advice of our panels, we followed Stineman et al.'s decision to begin with an initial partition of the data by Rehabilitation Impairment Category (RIC). Subsequent to the first publication on FIM-FRGs, Dr. Stineman expanded the 18 RICs to 20 by including a new RIC for Guillain-Barré Syndrome (previously included in neurological diseases), and splitting the RIC for major multiple trauma into two categories depending on the presence of injuries to the brain or spinal cord (Stineman et al., 1996). Sample sizes were too small to allow creation of separate RICs for burns, debilities, or developmental disabilities as our first technical advisory panel had wished. Therefore, these groups are treated together in a miscellaneous category (RIC 18, labeled "other" below).

Our empirical analysis sought to partition cases within each RIC to maximize explained resource use. In some RICs, we had enough cases to fit numerous parameters--possibly creating 20 or more groups within each large RIC. For administrative simplicity, we did not wish to create an excessive number of groups. In addition, consistent with the recommendation of our second panel, we did not want to create groups differentiated by very small intervals of motor or cognitive scales for fear it would encourage upcoding. It is not enough to simply fit the "best" model (one with maximal predictive power).

The original FRGs were developed using only the FIM motor score, the FIM cognitive score, and age. Readmission cases, defined as a second or later hospitalization for the same impairment, were omitted from those analyses. Interrupted stays were not tested for different resource needs. We analyze the most appropriate way of classifying both readmissions and interrupted stays.¹

Information about secondary conditions appears on the UDSmr, but codes are frequently invalid and, in a private communication, Dr.

¹The UDSmr contains questions that flag both types of cases.

Stineman told us that her exploratory analyses using these data suggested that they were of little help in explaining resource use. For this study, however, we have available the diagnoses appearing on the MEDPAR. Since these are typically coded and/or reviewed by trained medical record technicians (at least in PPS hospitals), it is probable that they are more accurate than the UDSmr information. An examination of the values entered in the diagnosis fields suggests that illegal codes are found much more frequently on the UDSmr than on the MEDPAR. Thus, we analyze whether the comorbidities described on the MEDPAR file can predict resource use and, if so, how this information should be used to classify patients.

Another issue is whether the classification model should be based on Medicare cases alone or should include data on cases from other payers. It is possible that the best classification model for Medicare cases would be obtained by restricting the data to Medicare cases alone. However, if a classification system fit all cases, it might be more easily accepted by hospital administrators because it could be used for a variety of management purposes. The original DRGs were defined on an all-payer dataset and have been used for many purposes beyond Medicare payment. Thus, as explained in Section 2, our Classification and Payment System Design Panel urged us to build the classification system using all UDSmr cases. The empirical issue we examine here is whether the use of all UDSmr cases results in a loss of predictive accuracy for Medicare cases.

Our strategy was to evaluate a series of alternative classification models based on their statistical power and the amount of bias in predictions for cases grouped by demographic and hospital characteristics. The models were fit to both Medicare data and all UDSmr cases. This strategy allows us to present the explicit trade-offs between statistical accuracy and other attributes--simplicity and use of non-Medicare data.

Our analyses used the UDSmr definition of a case--that is, interrupted stays are counted as a single case rather than as the multiple discharges found on the MEDPAR.

VARIABLE DEFINITIONS

Dependent Variables

Length of stay (LOS) was used as the measure of resource use in the original creation of the FRGs (Stineman et al., 1994). We used it as our only measure of resource use for all UDSmr cases. For interrupted-stay cases, LOS is defined as the total number of nights spent in the rehabilitation facility and is calculated by subtracting time spent in acute care. We also used LOS as the dependent variable in models based solely on the matched Medicare sample, in order to compare the resulting classification to that from the use of all UDSmr cases.

For the matched Medicare sample, we used an estimate of accounting cost as the dependent variable in another set of models. The cost estimate was derived from MEDPAR information about charges in each department, using the departmental cost-to-charge ratio (CCR) method (see, for example, Newhouse et al., 1989). Cost models could not be fit on the entire UDSmr sample because UDSmr does not contain departmental charges (although it contains total charges).² A comparison of the matched sample models based on LOS with those based on cost will show how sensitive the results are to the measure of resource use.

The CCRs were calculated from the PPS 10 Exempt Hospital and Excluded Unit file. When facility data from this file were not available, we used the PPS 10 Minimum Dataset. Because some departmental CCRs were beyond plausibility, values were trimmed to plus or minus three standard deviations of the mean. Separate limits were applied for freestanding hospitals and for units within acute-care hospitals. Cost reports were missing for seven hospitals. For these, median CCRs (within type of hospital) were used to calculate costs.

²We did not use this field because some, but not all, UDSmr facilities include physician charges as well as facility charges. This inconsistency introduces error in measuring resource use as well as a conceptual problem since the payment system that we are building will not cover physician charges. In addition, Dr. Stineman informed us that her analyses using charges had resulted in no improvement in FRG definition over using LOS.

For interrupted stays, the costs associated with each discharge were summed across MEDPAR records to provide an estimate of the final cost of the entire interrupted stay. .

Early results showed that the distribution of these variables were heavily skewed to the right. We looked at normal probability plots of raw data as well as residuals from our regression models, and saw that the logarithms symmetrized the residuals and made their tails look closely normal. Therefore, we use logarithms in all the classification work we describe.

Thus, we use three dependent variables, each transformed by taking the logarithm to the base 10, which we call:

- UDSLOS - LOS on UDSmr population
- MEDLOS - LOS on Medicare/UDSmr matched samples
- MEDCOST - Cost on Medicare/UDSmr matched samples.

Independent Variables

The independent variables that we tried as potential predictors in our models are defined in the first seven rows of Table 4.1 (i.e., AGE through POOR). The motor score is the sum of the score on 13 motor functional independence items; since each item score ranges from 1 to 7, the motor score may range from 13 to 91. The cognitive score is the sum of the remaining five functional independence items.

We did not use Rasch analysis to transform the FIM scales because such a transformation cannot affect a CART analysis. Rasch analysis results in a monotonic transformation of a scale and since CART searches for any partition of the scale, exactly the same partitions are searched with a monotonically transformed scale as with an untransformed scale.³

³The only exception to this would be if information is thrown away via rounding and thus the transformation becomes one which is not strictly monotonic. In this case, because of the missing information, CART can do worse on the Rasch transformed scale than on the untransformed scale. This is apparently what happened when Stineman et al. used the transformed Rasch scale in their analyses.

Table 4.1
Independent Variable Definition

Variable	Definition
AGE	age at admission (in years)
AMOT	FIM motor score: sum of functional independence items related to eating, grooming, bathing, dressing (2 items), toileting, bladder management, bowel management, and mobility (5 items)
ACOG	FIM cognitive score: sum of functional independence items related to comprehension, expression, social interaction, problem solving, and memory
INT_STAY	interrupted-stay indicator (1 = was an interrupted stay, 0 otherwise)
READMISSION	readmission (1 = readmission, 0 otherwise), from the UDSmr admission class item
COMORB	comorbidity (1 indicates the presence of one or more comorbidities) from the MEDPAR diagnoses. (See Appendix B.)
POOR	1 indicates either an SSI recipient (from the MEDPAR data) or Medicaid is a payer (based on UDSmr data)
FREESTANDING	binary indicator of type of facility
MARRIED	binary indicator based on UDSmr data
WHITE	binary indicator based on UDSmr data
MALE	binary indicator based on UDSmr data

In preliminary analyses, we explored the relationship between resource use and a variety of comorbidities suggested by our panels and the clinician authors. We ultimately decided on a binary variable indicating the presence of one or more of a selected set of comorbidities and complications. Appendix B describes the process used to define this comorbidity variable.

The last four variables in Table 4.1 were judged to be unsuitable for payment purposes. These variables were used only to examine the extent to which differences among groups and institutions remain in the data after accounting for the previous variables. If large differences in resource use existed by demographic categories, then hospitals might preferentially select patients with demographic characteristics associated with less-expensive stays.

SAMPLE SIZES

Table 4.2 shows the number of 1994 UDSmr cases in each RIC. For our analyses, certain cases were eliminated because their inclusion

Table 4.2
Sample Size by RIC

RIC	Label	Name	Total UDSmr	UDSLOS Analysis	MEDLOS Analysis	MEDCOST Analysis
01	STR	Stroke	43005	39844	28377	28322
02	BDT	Brain dysfunction, traumatic	5733	5042	1143	1134
03	BDN	Brain dysfunction, nontraumatic	4385	3888	1904	1898
04	SCT	Spinal cord dysfunction, traumatic	3291	3015	623	618
05	SCN	Spinal cord dysfunction, nontraumatic	4549	4094	2532	2521
06	NEU	Neurological conditions	7145	5966	4030	4014
07	OR1	Orthopedic--lower extremity fracture	17721	16681	13746	13762
08	OR2	Orthopedic--lower extremity joint repl	21945	20902	16956	16934
09	OR3	Orthopedic other	5543	5100	3878	3868
10	AM1	Amputation, lower extremity	1513	4915	3507	3495
11	AM2	Amputation, other	1715	365	259	258
12	AR1	Osteoarthritis	5455	1399	1199	1193
13	AR2	Rheumatoid and other arthritis	409	1570	1272	1270
14	CAR	Cardiac	2977	2654	2232	2220
15	PUL	Pulmonary	2632	2249	1787	1779
16	PAI	Pain syndrome	1980	1814	1161	1156
17	MNB	Major multiple trauma, without injury to brain or spinal cord	2015	986	323	321
18	OTH	Other disabling impairments	7347	6435	4910	4886
19	GB	Guillain-Barré	694	559	233	231
20	MWB	Major multiple trauma, with injury to brain or spinal cord	978	836	113	112
All	Total		139360	128344	90201	89976

would have distorted our results. For the analyses of UDSLOS, we eliminated children 16 years of age or younger, stays of less than three days, and cases with a LOS outside of three standard deviations of the mean of the log of the LOS. Because they likely did not receive a full course of rehabilitation, we also eliminated patients who died during their rehabilitation hospitalization, patients transferred to another rehabilitation hospital, and patients transferred to acute care who did

not qualify as an interrupted stay.⁴ The number of cases remaining in each RIC for our analysis with UDSLOS as the dependent variable is shown in the second column of Table 4.2.

The analysis of LOS for matched MEDPAR cases used the same restrictions as the UDSLOS analysis, but added the restriction that the case be in our matched MEDPAR-UDSmr dataset. For the cost analysis of the matched MEDPAR sample, we selected the same cases as for the MEDPAR LOS analysis and then applied the further restriction that charges had to be positive. These sample sizes are also shown in Table 4.2.

To fit our models, we split the sample in about half, reserving the other half of the dataset for a validation sample.

STATISTICAL MODELS

Most of the alternative models are based on the recursive partitioning algorithm called Classification and Regression Trees (CART) (Breiman et al., 1984). This algorithm was used in the initial development of the FRGs (Stineman et al., 1994).

The CART algorithm examines a set of input independent variables and searches for a series of binary partitions that maximizes the percentage of variance in the dependent variable that is explained (R-square). At each iteration, CART searches for the two-way split of any existing group that maximizes R-square.

We used the 10-fold cross-validation method to determine the optimal number of splits in the final classification tree. CART's method is to divide the data into ten mutually exclusive sets of equal size, chosen at random. For each set, a tree with k nodes is fit on the other 90 percent of the data, and the squared error of the predictions is computed and summed over sets. CART then chooses the k with the minimum sum of squares error, and it fits a tree on the entire dataset with k nodes. Because this method resulted in many splits in large RICs, we also report results based on a stopping rule of 5, 10, or 15 categories within each RIC. (Table 4.3 lists all the models that we report.) In all cases we required a minimum of 20 cases in each FRG.

⁴To determine death and discharge destination, we used the discharge setting variable from the UDSmr.

Since the CART analysis used only half of our total sample, this means at least 40 cases will be available for estimating weights for the smallest FRGs.

The version of CART that we used is contained in S-Plus, a popular statistical package among statisticians. This software has a feature that complicates description: when you ask it to fit a tree with a specified number of nodes, it takes that as strong guidance but may decide to drop a node if it adds nothing to R-square, or add a node if its contribution is substantial. In our discussion below, we refer to best five- or ten-node trees according to the request we made, but the actual number of nodes may vary slightly.

The CART models search for main effects and interactions, but produce mainly interactions because they work recursively within groups that have already been split. We were concerned that models dominated by interaction terms would not perform as well as models that sought to determine main effects, so we also compare some ordinary least squares regression (OLS) models and some combinations of OLS and CART analyses to the CART models.

One of these models is an OLS regression model with main effects terms covering quintile regions (within RIC) of FIM motor score, FIM cognitive score, and age, along with dummy variables for readmission, interrupted stay, and comorbidities. We view this model as a "gold standard" that shows how resource use actually varies with changes in each of the independent variables.

In our initial analyses, we simply gave CART a complete list of independent variables and asked it to make the selection. We saw, however, that the main variables it used were motor scores, cognitive scores, and age. Motor score was by far the most influential variable, occupying the top nodes of the trees in all cases. However the gold standard model showed that interrupted stays and comorbidities were also important predictors. The indicator for readmissions, on the other hand, was neither a consistent nor strong predictor. Other models showed that interrupted stays and comorbidities really were main effects rather than interactions. Since CART does not adequately handle main effects, our final model consists of a CART analysis based only on FIM

motor score, FIM cognitive score, and age, with main effects for interrupted stays and cases with a comorbidity. This model performs almost as well as the gold standard model and has substantially fewer parameters.

We organized our presentation in Section 5 around research questions. Table 4.3 summarizes the models in the order in which they are discussed and the reader may find it a useful reference.

Table 4.3

Description of Models Used in Analysis, by Research Question

Research Question	Method	CART Stopping Rule	CART Independent Variable	OLS Independent Variable
No. of FRGs	CART	10-fold rule	AMOT, ACOG, AGE,	None
No. of FRGs	CART	max (5 groups, 10-fold rule)	AMOT, ACOG, AGE	None
No. of FRGs	CART	max (10 groups, 10-fold rule)	AMOT, ACOG, AGE	None
No. of FRGs	CART	max (15 groups, 10-fold rule)	AMOT, ACOG, AGE	None
New variables	OLS	NA	None	Quintiles for AMOT, ACOG, AGE, plus COMORB, READMISSION, INT_STAY,
New variables	CART	5 groups	AMOT, ACOG, AGE, COMORB, READMISSION, INT_STAY	None
New variables	CART	10 groups	AMOT, ACOG, AGE, COMORB, READMISSION, INT_STAY	None
New variables	Mixed	5 groups	AMOT, ACOG, AGE	FRG, plus COMORB, READMISSION, INT_STAY plus interaction of FRG and INT_STAY
New variables	Mixed	5 groups	AMOT, ACOG, AGE	FRG, plus COMORB, READMISSION, INT_STAY plus interaction of FRG and comorbidity
Final model	Mixed	max (5 groups, 10-fold rule)	AMOT, ACOG, AGE	FRG, plus COMORB, INT_STAY

5. RESULTS

We begin by addressing two of our research questions: (1) how much explanatory power can we achieve by increasing the number of FRGs within each RIC? and (2) what is the best way to use information on interrupted stays, readmission, and comorbidities in building a classification model? The analyses presented in these first two subsections point toward a single methodology as superior for all three models--whether built on all UDSmr cases and LOS, Medicare cases and LOS, or Medicare cases and cost. The subsequent two subsections evaluate the chosen model for each of the three dependent variable and dataset combinations. In those subsections, we first evaluate the accuracy of the models' predictions for hospital and demographic groups. Then we examine how well the chosen models perform in predicting the half of our dataset that was held out for evaluation. The final subsection shows that the FRGs defined from all adults in the UDSmr dataset predict Medicare resource use just as well as the FRGs defined using only Medicare data. These analyses allow us to recommend a single case classification system for use in the model for a Rehabilitation Prospective Payment System.

HOW MANY FRGS?

The original set of FRGs contained one to five FRGs per RIC for a total of 53 FRGs. Our dataset contains more than four times as many cases as that which was used in creating the original FRGs. Therefore, we expected and found that we could create many more FRGs in some of the RICs using 1994 data.

Table 5.1 shows the number of groups created by applying the 10-fold cross validation method to our data. As in the creation of the original FRGs, the only independent variables in the model were age, FIM motor score, and FIM cognitive score. The number of data points is roughly half that of Table 4.2 because we reserved half of our dataset for evaluation of the model. The first two columns are based on using all adult UDSmr cases and log LOS as the dependent variable in the CART analysis. The remaining columns report the result of restricting the

Table 5.1
Sample Size and Number of Groups in Best CART Fit

RIC	UDSmr LOS		Medicare LOS		Medicare Cost	
	N	Groups	N	Groups	N	Groups
1	19935	26	14188	41	14161	25
2	2503	19	554	3	583	3
3	1972	6	955	3	947	4
4	1516	6	330	5	290	4
5	2063	16	1271	7	1258	8
6	2958	4	2017	4	2008	6
7	8320	17	6864	10	6891	7
8	10458	22	8437	24	8511	33
9	2563	3	1965	4	1907	3
10	2440	3	1748	5	1752	2
11	181	3	135	3	124	2
12	693	4	594	4	603	2
13	794	3	649	4	622	4
14	1310	4	1092	4	1135	5
15	1113	3	887	2	896	5
16	917	2	573	2	586	2
17	487	4	149	2	172	0
18	3243	4	2461	4	2440	4
19	300	5	113	2	119	3
20	421	6	68	0	44	0
Total	64187	160	45050	133	45049	122

sample to Medicare cases and show the difference between LOS and cost as the dependent variable.

The CART algorithm suggests fitting a total of 160 FRGs to our dataset of all UDSmr cases compared to the 53 fit in the initial research. Eight of the 20 RICs contain more than five FRGs, and five contain more than ten. The number of groups created using only Medicare cases is somewhat lower and distributed differently by RIC. The reduction in the number of FRGs is concentrated in RICs in which Medicare patients are a smaller than expected proportion of the population such as those involving trauma (2, 4, 17, and 19), other brain and spinal cord injury (3 and 5), and Guillain-Barré (20). In a few RICs, the number of FRGs is larger when the data are restricted to Medicare cases.

Increasing FRGs to the extent suggested by the results in Table 5.1 would add to the administrative complexity of a payment system. Consequently, we wish to learn how much explanatory power is gained as

the number of FRGs increases. Table 5.2 shows the R-square statistics from several alternative models based on FIM motor score, FIM cognitive score, and age as independent variables. The first column for each dependent variable provides the R-square for the best CART fit which has the number of groups shown in Table 5.1. The remaining columns show how explanatory power would be reduced if, instead of the number of groups from the best fit, we reduced the number of groups to 5, 10, or 15. Models in which the number of groups exceeds the number in the best fit are left blank. For example, in the full UDSmr dataset, the best fit for RIC 3 contains only six groups, and we do not report a 10 or 15 group model.

Going from the best fit down to a 15-group model reduces R-square by at most 1 percentage point. Going from 15 groups to 10 groups

Table 5.2
R-Square Statistics from CART Models, by RIC and Number of Groups
(in percentage)

RIC	UDSmr LOS				Medicare LOS				Medicare Cost			
	Best CART	5- group	10- group	15- group	Best CART	5- group	10- group	15- group	Best CART	5- group	10- group	15- group
1	29	25	27	28	26	22	24	25	26	22	24	25
2	49	44	47	49	18				19			
3	28	26			17				26	22	26	
4	37	36			30	30			29	29		
5	34	30	33	34	29	28			30	29		
6	12				11	10			13	12		
7	15	12	14	14	12	11	12		12	11		
8	20	16	18	19	19	16	18	19	20	15	18	19
9	9				8				7			
10	11				10	10			8			
11	35	28			24				15			
12	17				18	18			10			
13	19				21				18			
14	20	20			19				24	24		
15	12				9				12	12		
16	5				8				7			
17	30				21				25			
18	15				15				15			
19	38	38			24				43			
20	47	46			23				29			

typically causes a further 1 percentage point reduction in R-square. Thus, this analysis shows that there is little predictive power to be gained by going to more than ten nodes in any RIC. The drop from ten to five is slightly larger, typically two points, and it is three or more points in four of the 11 cases.

In addition to the desired simplicity of minimizing the number of FRGs, we did not want to create groups defined by very small intervals for motor and cognitive scales for fear it would encourage gaming and case-mix creep. Consequently, we examined the structure of the five and ten group FRGs for the large RICs. We found that several of the FRGs added when ten groups were permitted were based on cognitive or motor score intervals of only one or two points. We deemed these to be unsatisfactory and consequently decided that a five-group FRG model would be an acceptable compromise between predictive accuracy and administrative considerations.

If we choose to use the five-group tree whenever the best CART fit has more than five groups and the best CART fit otherwise, the resulting model would provide quite satisfactory explanatory power. The within-RIC explanatory power obtained is quite comparable to those within the Major Diagnostic Category (MDC) explanatory power of DRGs--even though we do not have costly surgical procedures as a grouping device and have many fewer FRGs per RIC than there are DRGs per MDC. Freeman et al. (1995) report within MDC R-squares for log of charges in the fifth revision of the DRGs. They range from 0.08 through 0.48, with a median of 0.21 and an interquartile range of 0.17 through 0.29. Using our Medicare cost dataset as most comparable to their dataset, we have findings that are only slightly worse: a range in R-square per RIC of 0.07 through 0.43, a median of 0.165, and an interquartile range of 0.12 through 0.24. The FRGs compare more favorably to the performance of DRGs on medical acute-care cases: across MDCs, Freeman et al. report a range from 0.03 through 0.26, a median of 0.08, and an interquartile range of 0.08 through 0.10.

MODELING READMISSIONS, INTERRUPTED STAYS, AND COMORBIDITIES

Readmissions and interrupted stays were not examined during the initial formulation of FRGs. Information from the MEDPAR on comorbidities and complications was not available either. Here we address how, if at all, this information should be used to improve our classification system. Because accurate data on comorbidities are found only in the MEDPAR, we will concentrate on the two models based solely on Medicare cases.

Based on research described in Appendix B, we defined a single dummy variable indicating the presence of at least one of the selected comorbidities or complications (CCs). These include all conditions - included as "Major CCs" in HCFA's DRG refinement project and additional conditions that our clinician collaborators suggested would be likely to increase the resources used during a rehabilitation hospitalization: deep vein thrombosis, infection, tracheostomy, and ventilator status.

We performed several additional preliminary analyses using dummy variables for readmissions, interrupted stays, and comorbidities. First we used OLS regression to examine the effect of these variables on resource use while controlling for age, FIM motor score, and FIM cognitive score. We also entered these same variables into our CART model using stopping rules of 5, 10, and 15 groups.¹

The OLS regressions were intended to help us understand how resource use varied with all the dependent variables. Separate regressions were run for each RIC. The independent variables were entered in the following form:

- (1) Five levels of age (quintile indicators, within RIC),
- (2) Five levels of motor scores (quintile indicators, within RIC),
- (3) Five levels of cognitive scores (quintile indicators, within RIC),
- (4) interrupted-stay indicator,

¹These CART models were run using the original set of 18 RICs. Because there are so few Medicare cases in RICs 19 and 20, we are sure that the Medicare results would not change if we redid them with the latest definition of RIC.

- (5) readmission indicator, and
- (6) comorbidity indicator.

The full set of regression results is shown in Appendix C. It is our best description of how resource use changes with these variables. In most RICs, resource use decreases monotonically with increasing motor score. The pattern of age and cognitive score is much less consistent. Cognitive effects are usually insignificant. Exceptions include RIC 2, where lower cognitive scores increase resource use, and RIC 1, where patients with the highest and lowest scores use less than average resources. The effect of age differs by RIC. In many RICs, younger patients receive more resources. However, in the orthopedic RICs (7 through 9) as well as RIC 16 (pain), it is the oldest patients who receive more resources.

The interrupted-stay coefficients were always positive and were highly significant in all RICs except RIC 17, 19, and 20, where the sample sizes are very small. The magnitude of the coefficients suggests that, after controlling for age, FIM, readmission, and comorbidity, an interrupted stay costs 40 to 60 percent (depending on RIC) more than a similar non-interrupted case. The presence of a comorbidity results in a positive increase in the costs of a Medicare case in almost all RICs. The correspondence between comorbidity and LOS was weaker. There was no consistent relationship between readmission and resource use. Sometimes, when the readmission coefficient was significant, it was quite small.

The CART models with all six independent variables did not reflect the findings from the regression. For example, despite the highly significant coefficient on interrupted stay in the model for RIC 1 ($t = 12.9$), it did not enter the CART models with five groups. In other RICs, when interrupted stay did enter the CART analysis, it split only a small part of the sample. In another example, the comorbidity variable appeared in only five RICs in the five-group model and nine RICs at the 10-group model despite the regressions showing a significant effect of comorbidity on cost in 11 RICs.

Residual analysis of these CART models showed that substantial biases remained in the prediction of resource use for cases with interrupted stays and comorbidities. We believe these biases arise from omitted variables, which in turn arise because CART really analyzes only interactions and these variables are truly main effects. To test whether interrupted stay was a main effect, we ran the following additional set of regressions. We created a five-node CART tree based solely on age, FIM motor, and FIM cognitive score. Then we regressed resource use on dummy variables for each FRG (except one), interrupted stay, and the interaction of interrupted stay and FRG. Of the 61 coefficients fitted for the Medicare cost data, we could measure, with $p < 0.05$, a difference between the effect of interrupted stay in a particular FRG and the RIC average in exactly two cases--less than would be expected by chance. For the Medicare LOS data, there were only five significant coefficients, again easily explainable by chance. (Detailed results not shown.)²

The CART results also suggested that readmission was not an important predictor of either cost or LOS.

THE CHOSEN MODEL AND ITS STRUCTURE

Because it was clear that interrupted stay and comorbidity had substantial effects on resource use that CART could not adequately capture, we decided on the following form of the model. We first defined FRGs based on a CART model fit using only FIM motor score, cognitive score, and age. Then an OLS model was used to estimate a multiplicative coefficient for interrupted stay and comorbidity after controlling for FRG.

The CART model chosen is the best CART fit with five or fewer FRGs per RIC. Appendix D contains figures that describe these CART FRGs for the full UDSmr dataset. They show that motor scores are always at the highest level of the tree. Cognitive scores appear at the second level only in RIC 12, but also appear in RICs 2, 7, 8, 19 and 20. Age appears only in RICs 1 and 3. The remaining 12 RICs have FRGs defined only by

²Detailed results are available from the authors. Similar analyses were used to demonstrate that comorbidity acted as a main effect rather than an interaction.

motor score. Motor score was also the dominant variable in defining FRGs for Medicare cases for both LOS and cost.

The classifications produce meaningful variations on the dependent variables. Table 5.3 shows mean length of stay and mean cost within FRGs produced by the CART classification scheme. Differences of means among FRGs are highly significant; standard errors are small relative to the differences between them. These differences are shown visually in Figures 5.1 through 5.3. One point easily seen by comparing Figure 5.1 with Figure 5.2 is that the UDSmr dataset contains cases that span a much greater range of resource use, particularly in RICs 2 and 4.

The effects of interrupted stay from our regressions that control for FRG and comorbidity are quite large. See Table 5.4, which presents the data only for the Medicare cases. The interrupted-stay variable has a significant effect on LOS in all RICs except RICs 17 and 20. It has a significant effect on cost in all RICs except 4, 16, 17, and 20. The dependent variables are measured in the logarithm to the base 10. To help interpret the importance of the coefficients, we include a column, labeled "% of increase," which gives the percentage by which resource use for interrupted stays is larger than the resource use for a case in the same FRG with the same comorbidity value. The magnitude of the interrupted-stay effect is quite large--typically indicating a LOS or cost increase of 40 to 60 percent.

Table 5.5 shows similar data about the coefficients for comorbidities and complications. The table shows that the effect of comorbidity varies by RIC. In RIC 1, cases with comorbidities actually have a shorter LOS than otherwise similar cases. In many RICs (5, 7, 8, 9, 13, 17, and 18), cases with comorbidities have a longer LOS. In other RICs, the effect on LOS, if any, is too weak to measure.

Another finding from Table 5.5 is that comorbidities have a greater positive effect on cost than they do on LOS. The coefficient on cost is statistically significant ($p < 0.01$) in several additional RICs (a total of 12 RICs for cost vs. 7 positive for LOS), and it is larger. We believe this is because patients with these comorbidities frequently

Table 5.3

Mean LOS by FRG

a. FRGs Produced by the Minimum Node UDSmr LOS Model

LOS				LOS				LOS			
FRG	N	Avg.	Std. Error	FRG	N	Avg.	Std. Error	FRG	N	Avg.	Std. Error
15	3090	14.2	0.1	61	898	24	0.4	143	565	16.3	0.3
14	4084	19.4	0.1	75	1409	13.8	0.2	142	237	20.4	0.6
13	4530	24.9	0.2	74	1049	15.1	0.2	141	92	27.2	1.1
12	2552	27.5	0.3	73	1661	17.7	0.2	153	689	16.8	0.3
11	5653	32.6	0.2	72	1391	18.5	0.2	152	241	22.2	0.8
25	314	13	0.4	71	2851	21.8	0.2	151	206	29.5	1.2
24	411	18.9	0.7	85	3445	10.5	0.1	162	660	15.3	0.3
23	842	24.2	0.5	84	2810	11.7	0.1	161	237	19.1	0.6
22	532	40.5	1	83	1549	13.8	0.2	174	202	14.1	0.6
21	440	68.1	2	82	1431	15.1	0.2	173	87	19.3	1.3
35	437	15.5	0.4	81	1209	18.8	0.2	172	153	24.4	1
34	516	20.2	0.5	93	1073	14.2	0.2	171	57	41.8	3.2
33	557	27.3	0.6	92	984	17.8	0.3	184	524	13.5	0.3
32	172	32	1.4	91	480	21.9	0.5	183	751	16.1	0.3
31	234	47.6	2	103	478	16.3	0.5	182	966	19.4	0.3
45	300	18.8	0.8	102	988	19.6	0.3	181	951	24.7	0.4
44	244	27.1	1.1	101	1009	25.3	0.4	195	21	9.9	1.4
43	464	39	1.1	113	39	11.7	1.1	194	66	18.1	1.8
42	237	52.5	2	112	65	19.5	1.6	193	46	24.4	1.8
41	254	74.9	2.4	111	80	26.1	1.7	192	100	33.6	1.9
55	551	14.4	0.4	124	237	13.3	0.5	191	56	47.3	3.9
54	427	19.4	0.6	123	197	16.8	0.6	205	136	17	1.1
53	471	24.8	0.7	122	102	17.5	0.8	204	59	20	1.3
52	457	32.8	0.9	121	170	23.7	0.9	203	43	27.7	1.3
51	125	47.1	2.2	133	83	10.2	0.6	202	108	35.4	2.2
64	614	14.4	0.4	132	466	15.1	0.4	201	69	69.1	4.4
63	675	17.4	0.3	131	227	23.8	0.8				
62	821	20.7	0.4	144	450	12.8	0.3				

Table 5.3--continued

b. FRGs Produced by the Minimum Node Medpar LOS Model

LOS				LOS				LOS			
FRG	N	Avg.	Std. Error	FRG	N	Avg.	Std. Error	FRG	N	Avg.	Std. Error
15	601	12	0.3	72	1141	18.6	0.2	143	505	16.1	0.3
14	1444	15.4	0.2	71	3079	20.9	0.2	142	198	20.5	0.6
13	2951	19.3	0.2	85	1110	9.8	0.1	141	75	26	1.2
12	2632	24.1	0.2	84	2606	11.3	0.1	152	552	17.1	0.4
11	6561	29.4	0.2	83	1603	13	0.1	151	348	24.9	0.8
23	187	18.1	0.7	82	1889	14.3	0.1	162	276	14.3	0.5
22	245	23.7	0.7	81	1311	18.2	0.2	161	312	18.4	0.5
21	157	37.3	1.8	94	402	13.7	0.3	172	64	15.2	0.9
33	265	15.5	0.5	93	337	15.8	0.4	171	110	23.3	1.1
32	190	19.9	0.7	92	629	17.6	0.3	184	372	12.8	0.3
31	494	27.7	0.7	91	545	21	0.5	183	594	16	0.3
45	76	17.4	1.4	105	28	11.2	1.2	182	735	19.2	0.4
44	65	24.1	2.1	104	26	16.3	1.9	181	748	23.8	0.4
43	62	24.7	1.4	103	631	18.8	0.4	192	58	21.9	1.6
42	50	36.5	2.5	102	324	20.8	0.5	191	62	41.2	3.2
41	40	56.9	6.3	101	750	25.3	0.4	201	45	26.7	2.3
55	149	12.5	0.6	113	23	11.6	1.4				
54	448	16.6	0.4	112	55	19.1	1.6				
53	316	24	0.8	111	46	26.1	2.3				
52	277	30.8	1.1	124	141	13	0.6				
51	71	42.1	2.4	123	147	16	0.6				
64	71	12.1	0.8	122	129	17.7	0.8				
63	395	15.5	0.4	121	188	22.8	0.8				
62	347	18.1	0.5	134	60	10.2	0.7				
61	1200	21.9	0.3	133	383	15.3	0.4				
75	971	13.7	0.2	132	75	20.7	1.2				
74	535	16.1	0.3	131	105	26.4	1.2				
73	1172	16.3	0.2	144	362	12.7	0.3				

Table 5.3--continued

c. FRGs Produced by the Minimum Node Medpar Cost Model

LOS				LOS				LOS			
FRG	N	Avg.	Std. Error	FRG	N	Avg.	Std. Error	FRG	N	Avg.	Std. Error
15	1093	9056	162	73	1140	11767	169	153	134	17132	845
14	3227	12360	111	72	673	11844	215	152	112	23049	1889
13	3292	16442	143	71	2404	14162	149	151	50	30289	2948
12	906	17937	334	85	2718	6731	65	162	275	9562	378
11	5643	22007	157	84	2052	7577	84	161	311	13182	461
23	174	12781	527	83	1192	8878	138	171	172	14374	815
22	254	16474	538	82	1543	9753	128	184	595	9261	218
21	155	28630	1525	81	1006	12829	207	183	587	11410	259
34	186	10614	441	93	738	9401	188	182	451	13721	408
33	267	13931	465	92	777	11527	219	181	807	17025	364
32	310	19100	544	91	392	14288	462	193	39	12225	1281
31	184	25441	1214	102	982	12835	252	192	53	21516	1319
44	131	13762	715	101	770	18068	395	191	27	39678	4417
43	54	17915	1920	112	23	7351	770	201	44	18842	1989
42	66	23334	1618	111	101	15302	1137				
41	39	45406	4951	122	359	9540	261				
55	349	9278	307	121	244	13934	558				
54	278	13180	573	134	60	7119	486				
53	284	16578	559	133	346	10903	348				
52	276	22354	783	132	112	14208	734				
51	71	29770	1852	131	104	18390	1007				
65	20	8923	1041	145	245	7278	277				
64	348	9340	279	144	265	8937	316				
63	461	11719	284	143	398	11379	316				
62	464	14028	320	142	152	14977	636				
61	715	16206	354	141	75	20968	1397				
75	1504	9009	118	155	447	12866	352				
74	1170	10318	163	154	153	15878	769				

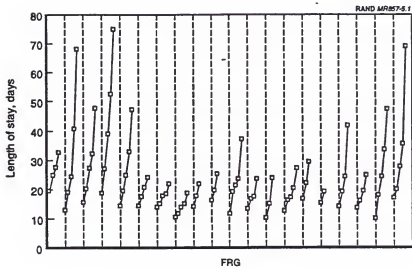


Figure 5.1-Average LOS in UDSmr FRGs, All UDSmr Cases

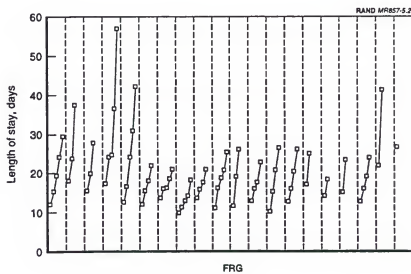


Figure 5.2-Average LOS in Medicare LOS FRGs, Medicare Cases

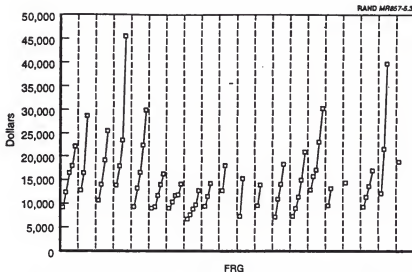


Figure 5.3-Average Cost in Medicare Cost FRGs,
Medicare Cases

Table 5.4

Coefficients on Interrupted Stays in Regression Models
Controlling for FRG and Comorbidities

RIC	Medpar LOS Model			Medicare Cost		
	Coeff.	% of Increase	t-stat	Coeff.	% of Increase	t-stat
1	0.1404	38	14.03	0.1406	38	12.99
2	0.2105	62	4.49	0.2265	68	4.53
3	0.1790	51	5.03	0.2037	60	5.46
4	0.1933	56	2.36	0.1676	47	1.86
5	0.2136	64	6.10	0.2117	63	5.62
6	0.1911	55	6.46	0.1985	58	6.27
7	0.1527	42	10.17	0.1519	42	9.21
8	0.1399	38	8.43	0.1503	41	8.28
9	0.1543	43	5.25	0.1758	50	5.64
10	0.1986	58	8.38	0.2541	80	9.38
11	0.3951	148	3.66	0.4591	188	3.90
12	0.2163	65	3.78	0.1836	53	3.04
13	0.1358	37	2.00	0.1521	42	2.11
14	0.0867	22	2.80	0.1259	34	3.45
15	0.1605	45	4.27	0.1242	33	2.81
16	0.1605	45	2.07	0.1691	48	1.79
17	0.0189	4	0.18	0.1565	43	1.23
18	0.1969	57	8.59	0.2236	67	8.79
19	0.2056	61	2.30	0.1890	55	2.06
20	0.1943	56	0.69	0.2344	72	0.81

require additional ancillary services leading to a higher cost per day as well as a longer LOS.³

The effect of comorbidities on cost per case is quite high in some RICs--it is 20 percent or higher in seven RICs--and yet in RIC 1 it is 0. Because of the large sample size in RIC 1, we are quite confident that the effect of comorbidity on cost is approximately 0--the increased cost per day roughly cancels out the decreased LOS. In other RICs, we cannot reliably measure any effect. We sought clinical advice about which comorbidity effects (and their observed direction) made sense, but little is understood about how comorbidities affect costs differentially across RICs. Ultimately, we settled on an empirical rule. If, in the

Table 5.5
Coefficients on Comorbidities in Regression Models
Controlling for FRG and Interrupted Stay

RIC	Medpar LOS Model			Medicare Cost Model		
	Coeff.	% of Increase	t-stat	Coeff.	% of Increase	t-stat
1	-0.0165	-4	-4.40	-0.0002	0	-0.05
2	0.0215	5	0.80	0.0254	6	0.88
3	0.0266	6	1.60	0.0524	13	3.00
4	0.0041	1	0.11	-0.0021	0	-0.05
5	0.0531	13	2.73	0.0893	23	4.24
6	0.0145	3	1.07	0.0478	12	3.31
7	0.0247	6	3.25	0.0643	16	7.70
8	0.0635	16	8.04	0.0824	21	9.53
9	0.0394	10	2.38	0.076	19	4.32
10	-0.0011	0	-0.09	0.0218	5	1.51
11	-0.0253	-6	-0.53	0.0246	6	0.47
12	0.0010	0	0.03	0.0509	12	1.60
13	0.0620	15	2.38	0.0924	24	3.32
14	0.0243	6	1.84	0.0795	20	5.06
15	0.0266	6	1.63	0.0403	10	2.11
16	0.0299	7	0.83	0.0938	24	2.13
17	0.1332	36	3.10	0.1922	56	3.58
18	0.0407	10	4.35	0.0603	15	5.82
19	0.0249	6	0.38	-0.0028	-1	-0.04
20	0.2306	70	1.90	0.1946	57	1.44

³The members of our second panel had hypothesized that, because of ancillary services, costs per day would be higher for patients with comorbidities. Dr. Elliot Roth was the first to express this opinion.

regression model containing the CART variables, the interrupted-stay variable, and the comorbidities variable, the comorbidity effect was positive with a t-statistic of two or more, we included it in the model. For the UDSmr LOS model, we did not include the comorbidity indicator, because it is not available for the population that cannot be matched to Medpar.

EVALUATION OF SELECTED MODEL STRUCTURE

Model Accuracy

The CART models explain variation in LOS and cost. Nevertheless, the simplifications made for the purpose of producing useful classification groups raise the possibility that they may lead to biased prediction rules. Therefore, we checked how well the FRGs predict actual log(LOS) or log(cost) for variables not included in the models: demographics and institutional type (see Section 4 for variable definitions).

For policy purposes, it is important that biases in prediction for demographic groups be small. From a hospital's perspective, demographics are a part of case mix. If demographics do not appear in the model, they should not be related to resource use. HCFA should also prefer that biases be small to reduce the incentives for skimming based on demographic characteristics. We predicted log(LOS) and log(cost) for each sample case in the final FRG model (based on either the best fit or five-nodes, whichever has fewer nodes). We then aggregated by hospital and case characteristics and calculated the percentage bias after exponentiation. Table 5.6 shows the subgroup percentage differences (relative to the mean) for LOS in the UDSmr population.

Previous studies have shown that age, gender, and marital status are predictors of the likelihood of going to an inpatient rehabilitation facility rather than to a skilled nursing facility (Kramer et al., 1994). However, our model is reasonably accurate for these demographic groups. The effects of age and gender are inconsistent across RICs. Married people typically use fewer resources than singles, probably because the availability of support in the home allows earlier

Table 5.6

Percentage of Errors in Minimum-Node CART Model

a. All UDSmr Cases LOS

Class	Age 65-75	Age 76-85	Female	Male	Not Married	Married	Not Poor	Poor	Non- White	White	Freestand. Hospital	Other Hospital
01	0	0	0	0	1	-1	0	-1	1	0	10	-6
02	-3	-13	-1	0	0	0	-2	9	-1	0	11	-7
03	1	-6	-1	1	1	-1	0	2	3	-1	21	-10
04	-18	-15	-5	2	3	-5	-2	6	4	-2	8	-6
05	-4	-3	-2	2	4	-4	-1	5	11	-2	10	-5
06	3	2	1	-2	4	-4	0	2	8	-1	8	-6
07	-3	2	0	-2	3	-6	-1	9	12	-1	9	-8
08	-2	5	2	-5	6	-8	0	7	13	-2	6	-5
09	-1	4	0	0	4	-6	0	2	11	-1	6	-7
10	4	7	4	-3	8	-7	0	-2	4	-1	9	-6
11	1	-7	-4	3	8	-9	0	0	19	-5	2	-2
12	-7	7	0	-1	4	-7	-2	22	9	-2	10	-15
13	1	1	0	-2	3	-4	-1	7	9	-1	5	-9
14	1	0	0	0	0	0	0	3	4	0	2	-2
15	2	1	2	-2	2	-3	0	2	-2	0	13	-19
16	-1	0	1	-1	1	-1	0	1	10	-1	1	-3
17	9	-1	1	-1	3	-4	0	3	7	-1	12	-6
18	0	-2	-1	1	2	-2	-1	7	11	-2	3	-4
19	-2	8	-3	3	6	-4	0	-1	-3	1	11	-5
20	-6	-11	-3	2	3	-6	-4	19	13	-3	11	-6
All	-1	1	0	-1	3	-4	0	4	6	-1	8	-6
N	19224	19943	37486	26671	35729	28428	58115	6042	10159	53998	29194	34963

Table 5.6-continued

b. Medicare Cases LOS

Class	Age 65-75	Age 76-85	Female	Male	Not Married	Married	Not Poor	Poor	Non- White	White	Freestand. Hospital	Other Hospital
01	1	1	0	0	1	1	0	0	2	0	11	-7
02	7	-6	4	-2	2	-1	0	-2	-1	0	14	-10
03	2	-2	0	0	2	-2	0	-6	-2	0	19	-9
04	2	5	-5	2	-1	-1	-1	12	0	0	11	-9
05	-2	0	0	0	4	1	1	-11	13	-2	10	-5
06	2	2	2	-3	3	-4	0	5	8	-1	10	-8
07	-5	1	0	-1	3	-4	-1	10	13	-1	10	-8
08	-4	4	2	-6	6	-6	0	9	15	-2	7	-6
09	-4	2	0	1	2	-8	0	4	12	-1	7	-9
10	1	4	4	-3	7	-4	0	2	3	-1	11	-7
11	5	-3	-1	1	4	-7	2	-	45	-6	2	-2
12	-9	5	1	-3	3	-6	-2	22	12	-2	11	-16
13	-1	0	0	1	1	-6	-1	15	5	-1	6	-13
14	0	0	0	0	0	-2	0	7	5	0	1	-1
15	2	0	2	-2	2	0	0	2	-6	0	13	-20
16	-2	-2	2	-5	2	-3	0	4	14	-1	2	-3
17	8	-5	4	-7	4	-3	-1	-	48	-4	10	-6
18	1	-1	1	-2	1	-7	0	4	11	-1	4	-5
19	-2	9	-3	3	-1	-2	1	-	-15	3	7	-4
20	6	-12	-3	3	20	1	-2	-	-	-1	6	-4
All	-1	1	1	-1	3	-4	0	4	7	-1	9	-7
N	17143	18429	28285	16866	26211	18940	42295	2856	5857	39294	20900	24251

Table 5.6-continued

c. Medicare Cases Cost

Class	Age 65-75	Age 76-85	Female	Male	Not Married	Married	Not Poor	Poor	Non- White	White	Freestand. Hospital	Other Hospital
01	2	0	-1	1	1	-1	0	0	4	-1	10	-6
02	7	-7	3	-2	4	-6	0	3	1	0	14	-11
03	2	-3	-3	2	0	0	1	-10	4	-1	17	-8
04	6	6	-2	1	-3	5	-1	7	-2	1	13	-10
05	-2	-1	-1	1	3	-3	1	-9	20	-3	7	-4
06	2	-1	2	-2	2	-3	0	4	13	-2	9	-7
07	-3	2	0	0	2	-4	-1	9	12	-1	8	-7
08	-4	3	2	-6	6	-8	-1	13	17	-2	5	-4
09	-4	2	-2	6	2	-3	0	0	16	-1	4	-6
10	2	0	4	-2	6	-6	0	2	8	-3	9	-6
11	2	-3	-4	2	3	-4	3	-	57	-8	2	-1
12	-6	5	0	0	1	-2	-1	12	6	-1	9	-13
13	1	-3	-1	3	-1	1	-1	14	9	-1	6	-12
14	1	-1	-1	1	-2	2	0	2	11	-1	-6	7
15	5	-4	0	0	1	-2	0	3	-3	0	13	-20
16	-2	-3	1	-2	2	-2	0	1	14	-1	-3	5
17	7	-3	3	-6	4	-6	-1	-	36	-3	19	-10
18	4	-3	1	-1	0	0	0	3	14	-2	4	-4
19	-2	10	-1	1	-2	1	1	-	-20	3	6	-4
20	4	-1	-2	2	15	-17	-2	-	-	-1	6	-4
All	0	1	0	-1	2	-3	0	4	9	-1	7	-6
N	17102	18398	28229	16820	26155	18894	42195	2854	5839	39210	20878	24171

discharge. We expect this is probably not a reasonable variable for a payment differential, and, in any case, the effect is quite small.

Those classified as "poor" (SSI recipients and Medicaid beneficiaries) typically use more resources than others, but in some RICs the effect is negative and in many others the effect was too small to measure well. As shown in our companion report (Carter et al., 1997), a hospital-level payment adjustment based on percentage of poor patients noticeably increases the fairness of payments because poor patients are concentrated in specific hospitals. The effect of nonwhite status on LOS is slightly larger than the effect of poverty, but nevertheless is still negative in several RICs. We expect that, as for acute care, the larger costs of poor patients and minority patients is traceable in part to care in expensive urban areas and much of this effect will be treated in other aspects of a rehabilitation PPS.

Table 5.6 also shows that freestanding rehabilitation hospitals use more resources than other hospitals with the same FRG case mix--most of which are exempt units of PPS hospitals. Again, this is an important point that is examined in greater detail in the companion report.

Explanatory Power

Table 5.7 summarizes the results of analyses for the entire final model--the regressions with dummies for each FRG (except one), interrupted stays, and comorbidities. The COMORB variable was used only in selected RICs in each Medicare model and was not available for the full dataset. The R-squares for the fitted model on the UDSmr dataset are typically only 1 to 3 percentage points higher than the R-square for the CART model alone, which were shown in Table 5.2. The improvements in the R-square for the Medicare sample are slightly higher than for the whole UDSmr sample. For example, four RICs for Medicare LOS and three RICs for cost had a 4 or more percentage point gain vs. none for the UDSmr sample. Despite the small gain in R-square for inclusion of the interrupted-stay and comorbidity coefficients, we believe the coefficients should be included in the classification model because, as shown in Tables 5.4 and 5.5, they identify cases with substantially larger resource requirements.

Table 5.7
Description of Model Structure and R-Squares
a. UDSmr LOS Model

RIC	No. of FRGs	CC in Model?	Fitted Sample		Evaluation Sample	
			n	R2	n	R2
1	5	no	19909	27	19935	27
2	5	no	2539	45	2503	41
3	5	no	1916	29	1972	26
4	5	no	1499	37	1516	31
5	5	no	2031	33	2063	30
6	4	no	3008	13	2958	15
7	5	no	8361	14	8320	13
8	5	no	10444	16	10458	16
9	3	no	2537	11	2563	10
10	3	no	2475	14	2440	14
11	3	no	184	27	181	15
12	4	no	706	19	693	17
13	3	no	776	20	794	11
14	4	no	1344	20	1310	20
15	3	no	1136	14	1113	16
16	2	no	897	6	917	5
17	4	no	499	31	487	23
18	4	no	3192	18	3243	16
19	5	no	289	39	300	38
20	5	no	415	46	421	35
total	82	na	64157	35	64187	34

b. MEDPAR LOS Model

1	5	no	14189	23	14188	24
2	3	no	589	21	554	14
3	3	no	949	20	955	15
4	5	no	293	31	330	8
5	5	yes	1261	30	1271	30
6	4	no	2013	12	2017	12
7	5	yes	6898	12	6864	11
8	5	yes	8519	17	8437	16
9	4	yes	1913	10	1965	10
10	5	no	1759	14	1748	13
11	3	no	124	29	135	8
12	4	no	605	19	594	15
13	4	yes	623	22	649	9
14	4	no	1140	19	1092	19
15	2	no	900	11	887	11
16	2	no	588	9	573	4
17	2	yes	174	18	149	10
18	4	yes	2449	19	2461	15
19	2	no	120	28	113	34
20	0	no	45	1	68	-2
total	71	na	45151	32	45050	31

Table 5.7--continued
c. MEDPAR Cost Model

RIC	No. of FRGs	CC in Model?	Fitted Sample		Evaluation Sample	
			n	R2	n	R2
1	5	no	14161	23	14161	24
2	3	no	583	22	551	12
3	4	yes	947	24	951	17
4	4	no	290	28	328	6
5	5	yes	1258	31	1263	30
6	5	yes	2008	15	2006	14
7	5	yes	6891	13	6855	12
8	5	yes	8511	17	8423	17
9	3	yes	1907	10	1961	11
10	2	no	1752	12	1743	12
11	2	no	124	24	134	7
12	2	no	603	12	590	11
13	4	yes	622	20	648	12
14	5	yes	1135	27	1085	26
15	5	yes	896	13	883	14
16	2	yes	586	8	570	5
17	0	yes	172	8	149	9
18	4	yes	2440	19	2446	16
19	3	no	119	40	112	23
20	0	no	44	1	68	0
total	68	na	45049	32	44927	31

Table 5.7 also shows what happened when we tested the models on the evaluation sample that was not used prior to selecting the model and was not used to fit the model. The column headed "Evaluation Sample R2" shows the percentage of variance explained by the prediction calculated from the fitting sample splits and regression weights. It shows the between-sample stability of predictions from the entire model, both structure and coefficients.

For the entire UDSLOS, the performance of the model on the evaluation sample is really quite good. Overall, the R-square changes very little: from 0.35 to 0.34. For six RICs, including the two largest (1 and 8), there is no deterioration in performance on the evaluation sample. There are only two RICs with a substantial decline in performance (a percentage decline greater than one-third)--RIC 11 with only 181 cases and RIC 13 with 794 cases.

For the Medicare samples, overall performance on the evaluation sample is also quite good. The R-square decreases from 0.30 to 0.29 for

the MEDPAR LOS model and from 0.32 to 0.31 for the MEDPAR cost model. However, more RICs have substantial declines in R-square. They tend to be smaller RICs. We will explore the reasons for these declines in more detail after we discuss the ability of the UDSmr FRGs to predict Medicare resource use.

FRG SELECTION

Our Classification and Payment System Design Panel recommended that we develop FRGs that would predict resource use for all patients, not just Medicare patients. We have already seen that the all-payer UDSmr sample identifies more FRGs and FRGs with greater resource requirements in some RICs, particularly trauma RICs and Guillain-Barré Syndrome, where the Medicare sample is an unusually small fraction of cases. Thus, the all-payer FRGs could be used for a variety of purposes where Medicare-based FRGs would be less useful. In this subsection, we ask how well the all-payer FRGs, developed from LOS on the entire UDSmr adult sample, predict LOS and cost for Medicare patients.

Medicare patients in each all-payer FRG within the same RIC use quite different levels of resources. Figure 5.5 shows mean LOS for Medicare cases in each all-payer FRG for both the fitting sample and the evaluation sample. Figure 5.6 shows mean cost per Medicare case for the same two samples. For completeness, Figure 5.4 provides similar LOS information for all adult UDSmr cases. Even in RICs with more all payer FRGs than Medicare FRGs (e.g., RICs 2, 19, and 20), the all-payer FRGs explain resource use for Medicare cases.

One other point can be seen in these figures. In the very high resource use FRGs (e.g., the highest resource use FRG in RICs 2, 4, 5, 17, 19, and 20), the fitting sample mean is substantially higher than the evaluation sample mean. This is partially a random phenomenon. These FRGs have a relatively small number of cases and the variance of resource use increases with the amount of resource use; the differences between the fitted and evaluation samples are not typically statistically significant. However, in addition to random variation, there is a regression to the mean phenomenon. In Figures 5.7 through 5.9, each data point is an FRG. These figures plot the average resource

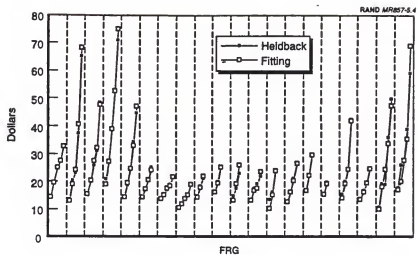


Figure 5.4-Average Cost for All UDSmr Cases in UDSmr FRGs, for Fitting and Evaluation Samples

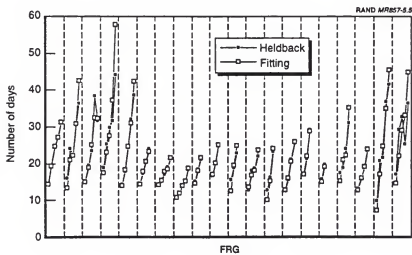


Figure 5.5-Average LOS for Medicare Cases in UDSmr FRGs, for Fitting and Evaluation Samples

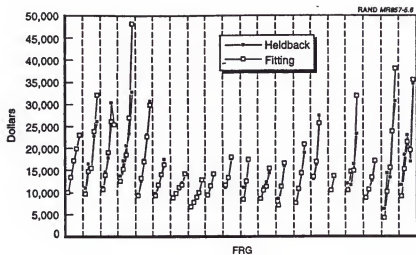


Figure 5.6—Average Cost of Medicare Cases in UDSmr FRGs, for Fitting and Evaluation Samples

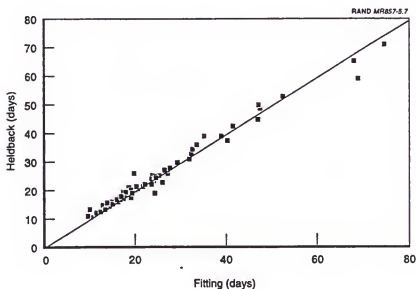


Figure 5.7—Scatter Plot of UDSmr FRG Values of Mean LOS in Evaluation Sample Against Mean LOS in Fitting Sample, All Cases

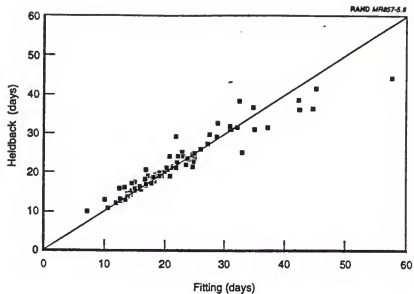


Figure 5.8-Scatter Plot of UDSmr FRG Values of Mean LOS in Evaluation Sample Against Mean LOS in Fitting Sample, Medicare Cases

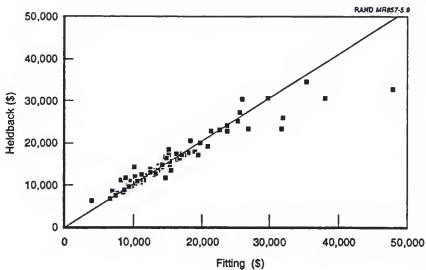


Figure 5.9-Scatter Plot of UDSmr FRG Values of Mean Cost in Evaluation Sample Against Mean Cost in Fitting Sample, Medicare Cases

use of cases in the evaluation sample against the average resource use of cases in the fitted sample. The straight line is at a 45-degree angle, and one would want the points to fall randomly along this line. There is a very good fit through the center, but there are several points (we judge two or three for the LOS curves and four for the cost curve) at the upper right where the fitted sample mean is substantially higher than the evaluation sample mean. A careful look shows the opposite phenomenon at the low end--the fitted mean is below the evaluation mean. This is the classic picture of regression to the mean--groups where cases are selected because they have a higher than average mean will contain higher than average errors, and so a replication of the selection rules will result in lower values for the highest original groups and vice versa. Although the validation sample has lower resource use than the fitting sample for the extreme groups in the upper right, it still has much larger than average resource use for these groups and so the group definitions remain quite reasonable despite this regression to the mean.

Table 5.8 compares the percentage of variance explained in a regression of LOS and cost in the Medicare sample on the all-payer FRGs with that from the FRGs defined based on the Medicare cases and the same dependent variable. The regressions all control for interrupted stay and comorbidities (in the same RICs as are shown in Table 5.7). Results are shown separately for the fitting and evaluation samples.

As Table 5.8 shows, we lose practically no explanatory power by using the all-payer FRGs from the UDS CART rather than Medicare-specific FRGs. In fact, the all-payer FRGs actually do slightly better than Medicare-specific FRGs in some RICs, perhaps because of clinically distinct categories that have a small number of Medicare cases that look more like similar non-Medicare cases than like other Medicare cases. Because these cases occur too infrequently in the Medicare data to allow the establishment of a separate FRG based on Medicare data alone, we classify these patients in a separate category only in the all-payer dataset.

Table 5.8

Effects of Using All-Payer FRGs on Medcost and MedLOS Datasets

RIC	Medcost						MedLOS					
	Fitting Sample			Heldback Sample			Fitting Sample			Heldback Sample		
	No. Obs	Medpar	UDS	No. Obs	Medpar	UDS	No. Obs	Medpar	UDS	No. Obs	Medpar	UDS
		CART	CART		CART	CART		CART	CART		CART	CART
01	14161	23	23	14161	24	24	14189	23	23	14188	24	24
02	583	22	23	551	12	15	589	21	23	554	14	16
03	947	24	24	951	17	17	949	20	21	955	15	17
04	290	28	28	328	6	4	293	31	27	330	8	7
05	1258	31	31	1263	30	31	1261	30	30	1271	30	29
06	2008	15	14	2006	14	15	2013	12	12	2017	12	13
07	6891	13	13	6855	12	13	6898	12	12	6864	11	12
08	8511	17	17	8423	17	17	8519	17	16	8437	16	16
09	1907	10	10	1961	11	11	1913	10	9	1965	10	9
10	1752	12	13	1743	12	14	1759	14	12	1748	13	13
11	124	24	26	134	7	14	124	29	28	135	8	7
12	603	12	16	590	11	15	605	19	19	594	15	17
13	622	20	19	648	12	11	623	22	21	649	9	10
14	1135	27	26	1085	26	25	1140	19	19	1092	19	19
15	896	13	11	883	14	15	900	11	12	887	11	13
16	586	8	6	570	5	6	588	9	7	573	4	5
17	172	8	25	149	9	5	174	18	22	149	10	11
18	2440	19	19	2446	16	15	2449	19	19	2461	15	15
19	119	40	43	112	23	15	120	28	33	113	34	28
20	44	1	39	68	0	24	45	1	34	68	-2	10

As we saw in Table 5.7, in Table 5.8 there are several RICs where the evaluation sample has a substantially smaller R-square than the fitting sample. We believe this results from the poor prediction of mean resource use for one FRG in these RICs. The FRGs with high cost and poor predictions in Figure 5.9 include one each from RICs 2, 4, and 17. Each of these poor predictions can be explained by either random variation or the regression to the mean we discussed earlier. The R-square measure is very sensitive to such extreme cases. Because the FRG definitions themselves remain robust, prediction can be improved by fitting the model to the entire sample and thus doubling the available sample. If HCFA chooses to implement an FRG-based PPS, it will have an even larger sample, and thus we believe that the apparent decline in R-square in a few RICs between fitting and evaluation samples should not be a serious threat to the accuracy of the payment formula.

We can use the data in Table 5.8 to compare the performance of our final model to that of the fifth revision of the DRGs. As discussed earlier, Freeman et al. (1995) report within MDC R-squares for log of charges that range from 0.08 through 0.48, with a median of 0.21 and an interquartile range of 0.17 through 0.29. Using the Medicare cost fitting data, we have similar findings: a range in R-square per RIC of 0.06 through 0.43, a median of 0.21, and an interquartile range of 0.14 through 0.26. Although we used different criteria to eliminate extreme cases than Freeman et al. did, we believe this has only a modest effect on the comparison. Freeman et al. report an overall R-square for all cases (including extreme cases) of 0.2814, again roughly comparable to our overall R-square of 0.2619 (which should be strictly comparable because it includes extreme cases, transfers, deaths, etc.).

6. CONCLUSIONS

FIM-FRGs defined by CART analysis within clinically homogeneous rehabilitation impairment categories (RIC) are robust, effective predictors of resource use. FRGs based on independent variables of motor FIM score, cognitive FIM score, and age were stable over time, as we demonstrated by using FRGs fit on 1990 data to predict 1994 data. Groups fit on all-payer UDSmr cases were as effective in predicting both the LOS and cost of Medicare cases as were groups defined to optimize fit to these data. For administrative simplicity and to reduce the number of cases where small changes in score would cause changes in reimbursement, we recommend that the number of FRGs per RIC be limited to a maximum of 5.

Although CART analysis defines FRGs that do predict resource use, it does not adequately capture main effects. We find that cases with interrupted stays and, in many RICs, comorbidities or complications cost substantially more than other cases in the same FRG. Thus, we suggest using a RIC-specific multiplier for the weight of such cases. Readmissions were found to differ little from other cases in the same FRG.

The explanatory power at the case level of the FRGs with comorbidity and interrupted-stay effects is roughly equal to that of the fifth revision of the DRGs. The median within RIC explanatory power of the FRGs is 0.21, which equals the median within MDC explanatory power of the DRGs. The explanatory power exceeds that of DRGs for medical cases (median within MDC explanatory power is 0.08). In assessing the adequacy of this degree of explanatory power, it is important to realize that other elements of a rehabilitation PPS such as outlier payments and facility adjustments for area wage rates and disproportionate share will increase the percentage of variance in costs at the hospital level that are explained by reimbursement under an RPPS. For example, in the acute PPS, the correlation coefficient between costs and reimbursements at the hospital level was reported at 0.82, with an implied R-square of 0.67

(Carter and Rumpel, 1993). Carter et al. (1997) provide similar information for alternative specifications of an RPPS.

Our definition of the comorbidity variable was derived empirically. It would be worthwhile for HCFA to examine whether it is appropriate to pay for all the conditions identified here as comorbidities. For example, HCFA might restrict ventilator cases to those on ventilator at admission to rehabilitation. Some infection conditions may be judged to be caused by poor quality of care and thus not be appropriate for additional payment. We expect that even after making such changes there will be a substantial relationship between the remaining comorbidities and cost per case.

The analysis in this report has several implications for the model rehabilitation PPS we are building. In Carter et al. (1997), we analyze simulated payments to show that facility-level adjusters can substantially ameliorate biases in the FRGs against adequate payment for poor patients.

Our advisory panels generally endorsed development of a rehabilitation PPS as a method of improving the undesirable incentives that exist in the current TEFRA payment system. But they cautioned that there was a need to independently verify the reliability of FIM scoring. The UDSmr is preparing additional training materials, and the effect of these training materials on reliability should be evaluated. Our panel also suggested that certain cognitive areas, currently unmeasured by the FIM, might be important additional predictors of resource use. They also recommended clarification of rules for assignment of impairment code.

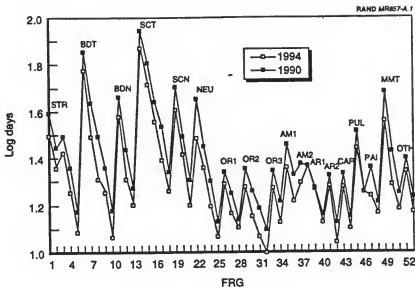
Appendix A
REPLICATION OF 1990 FRGs

In preliminary work, using a sample different from that used in the main body of this report, we addressed the stability of the FRGs. We examined the extent to which the original FRGs defined using 1990 UDSmr data in Stineman et al. (1994) were able to explain resource use during 1994. These FRGs were based on applying CART to the log of LOS using FIM motor score, FIM cognitive score, and age as the independent variables. The 1994 data are much more extensive than the 1990 data (139,000 cases from 352 hospitals vs. 54,000 from 125 hospitals). Consequently, we also examined whether replicating the original methodology on this larger dataset would lead to similar definitions of FRGs.

The original research effort by Stineman et al. excluded cases classified as for evaluation only and readmissions in order to get a more homogeneous dataset. For this replication analysis we also excluded such cases.¹ We followed Stineman in excluding cases of patients under 16, extremely long and short LOS cases, transfers, and deaths. These exclusion criteria left a study sample of 36,980 patients in 1990 and a replication study sample of 116,010 patients in 1994.

We began by classifying each patient in our 1994 replication study sample into one of the FRGs defined in Stineman et al. (1994). Figure A.1 compares the mean LOS in 1994 in each FRG (triangles) with the 1990 mean LOS reported by Stineman et al. (circles). LOS decreased between 1994 and 1990, as one can see (the triangle is always below the circle for the same FRG). However, 1994 LOS is still well explained by the 1990 FRGs--the 1994 LOS varies across the FRGs in each RIC. In fact, the 1994 mean LOS seem to be a constant amount below the 1990 mean LOS for the same FRG. Because the plot is in log scale, the constant

¹These cases are included in the analyses described in the main body of the report.



**Figure A.1—Comparison of Average LOS in 1994
with Average LOS in 1990, by 1990 FRG**

difference between 1990 and 1994 data means that LOS declined by the same proportion in each FRG.

The first column of Table A.1 presents the published data on the percentage of variance in 1990 LOS explained by the FRGs. There are only 18 RICs because RIC 20 (Guillain-Barré Syndrome) was included within RIC 06 (neurological), and RIC 19 (major multiple trauma with brain or spinal cord injury) was included in RIC 17 (major multiple trauma). The second column compares the percentage of variance in 1994 LOS explained by the 1990 FRGs. Although the 1994 data were not used to define these FRGs, the FRGs explain the 1994 data just as well as they explain the original 1990 data.

The remaining question we ask is the extent to which the 1994 data would create substantially different FRGs. We used the CART algorithm to create new FRGs using the 1994 replication study sample. Because the large sample size would virtually guarantee that we could find additional statistically significant splits, we restricted the CART algorithm to producing exactly the same number of splits as were produced in the original. The last column of Table A.1 shows that these

Table A.1
Replication Results: R-Squared

RIC	1990 Data ^a	1990 FRGs Used to Predict to 1994 Data	1994 FRG Replication
01 STR	23	25	27
02 BDT	39	43	45
03 BDN	13	20	22
04 SCT	30	39	40
05 SCN	17	26	26
06 NEU	14	14	16
07 OR1	9	7	10
08 OR2	14	14	14
09 OR3	10	6	7
10 AM1	7	6	9
11 AM2	0	0	17
12 AR1	10	17	16
13 AR2	12	7	13
14 CAR	6	11	14
15 PUL	24	10	11
16 PAI	2	2	5
17 MMT	32	24	33
18 OTH	11	10	11
Overall	31	33	34

^aFrom Stineman et al., 1994.

new FRGs explain only slightly more of the variance in 1994 LOS than the 1990 FRGs.

The general structure of the FRGs in the replication was quite similar to the original, although individual cut points differ between the two years. The motor score was the source of the first split and the predominant predictor in all RICs in both years. Age entered only in RIC 1 in the original FRGs, but in both RIC 1 and RIC 4 in the replication. The cognitive score entered in four RICs in the original FRGs (2, 3, 12, and 17) but in only two RICs in the replication (2 and 12).

Appendix B

COMORBIDITIES

Appendix B records the process used to define the final comorbidity variable used in our classification work. We began by using clinical judgment. Drs. Rosenfeld and Rubenstein defined a set of comorbidity groupings that they hypothesized might increase resources used by inpatient rehabilitation cases. We subsequently tested each of these variables and developed a single binary indicator for the presence of any of a selected set of comorbidities.

COMORBIDITY GROUPINGS

The comorbidity groupings that we tested in preliminary work are defined in Table B.1. In addition to the variables in the table, we tested a variable defined as 1 if and only if the patient has a "major CC" according to the definition in HCFA's new refined DRGs (*Federal Register*, May 27, 1994).

Some of the variables in Table B.1 are collinear, and thus cannot be entered together (e.g., all psychiatric disorders with the four individual psychiatric categories). A few of the variables depend on RIC as well as comorbidities. For example, the pulmonary variable will be a comorbidity only for cases not in the pulmonary RIC. In addition, we considered an "any CC" category that would be defined as 1 for a patient having any one of the variables in the table (or perhaps any one except those shown individually to have no effect). A few comments on the table groupings follow:

End Stage Renal Disease (ESRD)

Our working hypothesis was that patients on dialysis are more costly. The variable ESRD2 may identify some patients who are on dialysis for which the V451 was not coded. However, it will also identify patients with lesser amounts of renal insufficiency. These patients may also require additional costs, but one would expect these costs to be lower than for dialysis patients.

Infections

Infections require resources for diagnosis, they may delay treatment, and the delay may result in deconditioning. Urinary tract infection, which is not included, may sometimes have similar effects. We eliminated it from consideration for payment because it covers a wide spectrum of illness and is therefore gameable. Also, it may represent the result of care rather than being exogenous. It would be desirable to identify persons with an indwelling Foley catheter, but the ICD-9-CM diagnosis system does not allow identification of such patients.

Tracheostomy

This may be a marker of a severely ill patient rather than an actual cause of extra care.

Lung Diseases

The four lung variables will be zero for cases in the pulmonary RIC. We also should determine whether the ventilator variable interacts with the pulmonary RIC.

Most of the remaining variables are individual conditions selected by the physicians. The heart variables (CHF, ATRF, ARRHYT) are zero for cases in the cardiac RIC. The AMP variable is zero for cases in the relevant amputation RIC.

We defined the variables only for cases found on the MEDPAR because the diagnosis fields are much cleaner on this file (at least in the formal sense of legal codes).

ANALYSES

We began by entering each of the comorbidity variables into a regression of the log of LOS, which controlled for FIM and age. Regressions were run separately for each of 17 RICs defined by Stineman et al. in their original work (excluding "other"). A significance level of 0.05 is used below unless otherwise discussed. The results were as follows:

1. The set of comorbidities defined as a "major CC" according to the definition in HCFA's new refined DRGs was usually positive and was significant in four of the 17 RICs tried.

2. The psychiatric variables, when significant, were most likely to be negatively related to resource use (out of the $17 * 4 = 68$ individual variable coefficients, 12 were negative and significant and only two were positive and significant).

3. The variables INFECT, DVT, TRACH, and VENT were almost always positive, and were statistically significant in several cases (counts of significant coefficients were five, six, four, and two, respectively, out of 17).

4. The malnutrition variable was almost always close to 0 (i.e., insignificant at the 0.20 level in 12 of the 17 regressions). It was significant (i.e., $p = 0.05$ as before) and negative in one regression and significant and positive in two regressions.

5. None of the other variables was positive and significant in more than one out of 17 RICs and none was dominantly positive even at the $p = 0.2$ level.

Final Comorbidity Variable Definition

Based on this first set of regressions, we created a new comorbidity variable for patients who had a "major CC," or INFECT, DVT, TRACH, and VENT. We used this new variable to identify patients with a relevant comorbidity in all further work.

Discussion

The result of our analysis is a single comorbidity variable that can be used to split FRGs in RICs where comorbidity has a significant affect on resource use. This is analogous to the "complication and comorbidity (cc)" variable used to split DRGs in the original PPS DRGs and which is still the basis of the DRGs used for Medicare payment in the acute PPS.

Freeman et al. (1995) developed a more sophisticated split of comorbidities. The first steps of their procedure resembled our steps. They began with groups of ICD-9-CM that measured clinically meaningful comorbidity categories. These DXCATs are similar, at least in intent, to the comorbidity groupings in Table B.1. They then used regression analyses to determine which of the DXCATs were associated with significantly longer LOS and higher charges. An additional step divided

the significant DXCATs into classes based on the magnitude by which the significant DXCATs increased resource use. There were three classes for medical hospitalizations and four classes for surgical DRGs.

Our data are not capable of measuring the effects of comorbidities precisely enough to implement a comorbidity grouping like that of Freeman et al. The primary reason is that we have a much smaller sample--fewer than 100,000 Medicare cases (the source of our comorbidity information). Freeman et al. (1995) had 3.5 million cases from California plus over 600,000 cases from Maryland, plus almost 2 million Medicare discharges. In addition to the small sample, our data are likely to contain cases with uncoded comorbidities (see Carter et al., 1997).

Should HCFA choose to implement a RPPS, we expect comorbidity coding would improve. Further, a much larger sample would be available. Consequently, it would be appropriate to examine whether the comorbidity variable definition used here could be improved using data gathered from an operating system.

Table B.1
COMORBIDITY VARIABLES

Variable	Comments	ICD-9-CM Codes
ESRD1	Patients on dialysis	V451 Renal dialysis status
ESRD2	Patients may be on dialysis	V451 Renal dialysis status 585 Chronic renal failure 403.01 Malignant hypertensive renal disease, with renal failure 403.11 Benign hypertensive renal disease, with renal failure 403.91 Hypertensive renal disease unsp., with renal failure
INFECT	Selected infections	038.0-038.9, 054.5, 998.5, 098.89, 027.0 to 027.9 Septicemia 569.61, 380.10, 611.0, 607.2, 998.5, 373.13, 681.00 to 682.9, 608.4, 376.01, 597.0, 567.2, 608.0, 608.4, 528.3, 527.3 Cellulitis 707.0, 707.1, 440.23, 440.24 Decubitus ulcer 480.0 to 487.0, 514, 507.0 to 507.8, 518.3, 466.1, 515, 506.0, 112.4, 136.3, 513.0 Pneumonia 996.61 Infection shunt 996.64 *Infection indwelling urinary catheter 998.5 Post op wound infection
TRACH	Tracheostomy	V44.0, V55.0 Tracheostomy status
VENT	Ventilator status	V46.1 Respirator
LUNG1	Selected lung-related diseases	490 Bronchitis, not specified as acute or chronic 491x Chronic bronchitis 492x Emphysema 494xx Bronchiectasis 496 Chronic airway obstruction, NEC ^a 501 Abestosis 518xx Other diseases of lung 519x Other diseases of respiratory system
ASTHMA	Asthma	493xx Asthma
LUNGHRT	Chronic pulmonary heart disease	416x Chronic pulmonary heart disease
LUNG2	All lung-related diseases	Any condition listed in LUNG1, ASTHMA, or LUNGHRT
DEPRESS	Depressions	311, 300.4, 290.21, 298.0, 296.20 to 296.89,
DELERIM	Delirium	780.02, 780.09, 780.3, 291.0 to 291.8, 292.0 to 292.9, 293.89, 296.1, 296.0, 300.11
DEMENTIA	Dementias	290.0 to 290.43, 294.1, 295.00 to 295.15, 295.30 to 295.35, 094.1
PSYCH1	Selected psychiatric disorders	Codes 290 to 319, except those listed in DEPRESS, DELERIM, and DEMENTIA

Table B.1-continued

Variable	Comments	ICD-9-CM Codes
PSYCH2	All psychiatric disorders	Codes 290 to 319
DIAB1	Insulin-dependent diabetes (not out of control)	250.01, 250.11, 250.21, 250.31, 250.41, 250.51, 250.61, 250.71, 250.81, 250.91
DIAB2	Diabetes (not out of control)	250.00, 250.10, 250.20, 250.30, 250.40, 250.50, 250.60, 250.70, 250.80, 250.90
BRCANC	Brain cancer	191x Brain cancer
GIHEM	GI hemorrhage	578x Gastrointestinal hemorrhage
V730		730x Osteomyelitis, periostitis, and other infections of bone
AMP	Leg amputation above foot	V49.75, V49.76, V49.77 Amputation status
MALNUT	Malnutrition	260.0 to 263.9; 579.3
DVT	Deep venous thrombosis	451.19; 451.0 to 453.9
CHF	Coronary heart failure	428.0, 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93
ATRF	Atrial fibrillation	427.31, 427.3
ARRHYT	Other arrhythmias	427.0 to 427.2; 427.4 to 427.9

^aNot elsewhere classified.

Appendix C

GOLD STANDARD MODEL COEFFICIENTS AND T STATISTICS

Table C.1 displays the regression results from the main effects models. Regressions are provided for each of the three modeling datasets we used: MEDPAR cost, MEDPAR length of stay, and UDSmr length of stay. Within each dataset, a regression is provided for each of the 20 RICs.

The dependent variable is measured in base 10 logarithms

The independent variables are all dichotomous.

Variable	Description
acog1	cognitive scores--quintile indicators, within RIC (lowest to highest); acog5 is always the omitted category
acog2	
acog3	
acog4	
agef1	age--quintile indicators, within RIC (lowest to highest); agef5 is always the omitted category
agef2	
agef3	
agef4	
amot1	motor scores--quintile indicators, within RIC (lowest to highest); amot5 is always the omitted category
amot2	
amot3	
amot4	
int.stay	interrupted stay indicator
comorb	comorbidity indicator
readmit	readmission indicator

Table C.2 presents the T statistics from the gold standard models.

Table C.1
Regression Coefficients from Gold Standard Models

	RIC									
	01	02	03	04	05	06	07	08	09	10
MEDPAR Cost										
constant	3.89	3.89	3.89	4.03	3.83	3.87	3.86	3.77	3.86	3.95
acog1	0.02	0.14	0.07	-0.04	0.04	-0.01	0.04	0.09	0.03	0.04
acog2	0.03	0.08	0.06	0.01	0.05	0.01	0.06	0.04	0.04	0.02
acog3	0.04	0.07	0.05	-0.03	0.03	-0.01	0.04	0.00	0.04	0.02
acog4	0.03	0.05	0.05	0.00	0.00	0.01	0.00	0.00	0.00	0.00
agef1	0.07	0.14	0.04	0.01	0.04	0.05	0.00	-0.05	0.01	0.01
agef2	0.07	0.08	0.06	0.04	0.02	0.05	0.00	-0.04	0.00	0.03
agef3	0.05	0.05	0.04	0.03	0.02	0.03	0.02	-0.05	0.02	0.03
agef4	0.04	0.05	0.02	0.03	0.03	0.03	0.02	-0.02	0.02	0.04
amot1	0.32	0.30	0.30	0.44	0.37	0.23	0.17	0.21	0.17	0.19
amot2	0.29	0.20	0.25	0.22	0.30	0.21	0.16	0.12	0.11	0.16
amot3	0.23	0.14	0.18	0.16	0.21	0.17	0.10	0.08	0.09	0.10
amot4	0.14	0.10	0.07	0.07	0.09	0.11	0.06	0.05	0.03	0.04
int.stay	0.14	0.24	0.22	0.21	0.22	0.20	0.15	0.15	0.17	0.26
comorb	0.00	0.05	0.05	0.02	0.09	0.05	0.06	0.08	0.07	0.01
readmit	-0.02	-0.05	-0.09	-0.14	0.01	-0.01	0.03	0.00	0.03	-0.04
MEDPAR LOS										
constant	1.08	1.05	1.08	1.19	1.03	1.07	1.10	0.99	1.07	1.18
acog1	-0.01	0.14	0.07	-0.06	0.04	-0.01	0.02	0.08	0.02	0.01
acog2	0.01	0.08	0.04	0.02	0.03	0.01	0.05	0.04	0.04	0.02
acog3	0.02	0.07	0.04	-0.01	0.03	-0.01	0.04	0.00	0.05	0.01
acog4	0.02	0.06	0.04	0.00	0.00	0.01	0.00	0.00	0.00	0.00
agef1	0.05	0.12	-0.01	0.00	0.03	0.02	-0.03	-0.06	-0.01	-0.04
agef2	0.05	0.06	0.04	0.02	0.00	0.03	-0.02	-0.05	-0.01	0.00
agef3	0.04	0.05	0.02	0.00	0.02	0.02	0.01	-0.05	0.00	-0.01
agef4	0.03	0.05	-0.02	0.02	0.03	0.03	0.01	-0.02	0.01	0.01
amot1	0.30	0.27	0.26	0.41	0.34	0.20	0.15	0.19	0.16	0.16
amot2	0.29	0.17	0.22	0.22	0.28	0.19	0.14	0.11	0.12	0.16
amot3	0.23	0.15	0.15	0.19	0.19	0.16	0.09	0.07	0.08	0.10
amot4	0.14	0.10	0.05	0.10	0.07	0.11	0.06	0.04	0.03	0.07
int.stay	0.14	0.23	0.18	0.26	0.22	0.19	0.15	0.15	0.15	0.21
comorb	-0.02	0.04	0.02	0.03	0.05	0.01	0.02	0.06	0.04	0.00
readmit	-0.01	-0.06	-0.05	-0.13	0.04	-0.01	0.02	0.01	0.04	-0.04

Table C.1-continued

	RIC									
	01	02	03	04	05	06	07	08	09	10
<u>UDSmr LOS</u>										
constant	1.06	0.98	1.02	1.11	1.03	1.07	1.09	0.99	1.05	1.15
acog1	0.01	0.19	0.10	0.02	0.03	0.00	0.02	0.07	0.03	0.03
acog2	0.02	0.15	0.08	0.02	0.03	0.01	0.05	0.03	0.04	0.03
acog3	0.03	0.11	0.05	0.00	0.04	0.02	0.04	0.00	0.04	0.02
acog4	0.02	0.07	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00
agef1	0.05	0.06	0.07	0.13	0.06	0.00	-0.04	-0.08	-0.03	-0.05
agef2	0.05	0.07	0.07	0.09	0.05	0.02	-0.02	-0.05	-0.02	-0.04
agef3	0.04	0.09	0.05	0.09	0.00	0.02	0.00	-0.04	0.00	-0.01
agef4	0.03	0.06	0.03	0.07	0.01	0.02	0.01	-0.03	0.02	0.01
amot1	0.34	0.51	0.35	0.57	0.39	0.22	0.17	0.19	0.18	0.19
amot2	0.31	0.34	0.26	0.40	0.30	0.19	0.15	0.11	0.12	0.17
amot3	0.25	0.21	0.17	0.33	0.20	0.16	0.09	0.07	0.09	0.09
amot4	0.14	0.11	0.09	0.17	0.10	0.09	0.06	0.05	0.05	0.08
int.stay	0.15	0.16	0.20	0.16	0.23	0.19	0.17	0.14	0.17	0.19
comorb	-0.02	-0.01	-0.01	-0.06	0.04	0.01	0.02	0.06	0.04	0.00
readmit	0.00	0.01	-0.06	-0.10	0.02	0.00	0.03	0.01	0.04	-0.03

	RIC									
	11	12	13	14	15	16	17	18	19	20
<u>MEDPAR Cost</u>										
constant	3.76	3.81	3.83	3.78	3.98	3.91	3.79	3.83	3.84	3.72
acog1	0.08	0.10	0.05	0.03	0.03	0.08	0.10	0.02	0.08	0.11
acog2	-0.08	0.11	0.06	0.02	0.02	0.10	0.08	0.02	0.06	0.20
acog3	-0.05	0.04	0.03	0.04	-0.01	0.03	0.09	0.02	0.00	0.03
acog4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08
agef1	0.08	0.03	0.04	0.00	0.07	-0.01	0.13	0.08	0.14	0.15
agef2	0.07	0.02	0.01	0.00	0.07	-0.02	0.08	0.06	0.00	0.01
agef3	0.07	0.06	0.00	-0.03	0.03	-0.02	-0.01	0.04	-0.02	0.17
agef4	0.14	0.05	0.00	-0.01	0.00	-0.06	0.05	0.02	0.07	-0.03
amot1	0.29	0.19	0.27	0.31	0.21	0.12	0.30	0.26	0.57	0.43
amot2	0.31	0.11	0.18	0.18	0.11	0.09	0.15	0.20	0.40	0.34
amot3	0.20	0.07	0.11	0.14	0.05	0.02	0.17	0.15	0.26	0.29
amot4	0.19	0.01	0.11	0.05	-0.01	-0.02	0.05	0.08	0.10	0.32
int.stay	0.46	0.17	0.15	0.13	0.12	0.18	0.10	0.22	0.23	0.13
comorb	0.01	0.05	0.10	0.08	0.04	0.08	0.16	0.06	-0.02	0.20
readmit	-0.08	0.04	-0.04	0.04	0.09	0.00	0.25	-0.02	0.06	0.45

Table C.1-continued

	RIC									
	11	12	13	14	15	16	17	18	19	20
<u>MEDPAR LOS</u>										
constant	1.00	1.03	1.00	1.04	1.14	1.10	1.04	1.04	1.10	0.79
acog1	0.03	0.10	0.04	0.02	0.01	0.06	0.05	0.00	0.06	0.13
acog2	-0.06	0.12	0.05	0.03	0.00	0.09	0.03	0.00	0.07	0.26
acog3	-0.01	0.06	0.02	0.04	-0.01	0.05	0.05	0.01	0.00	0.13
acog4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19
agef1	-0.03	-0.03	0.02	0.00	0.02	-0.04	0.05	0.04	0.09	0.20
agef2	0.02	-0.02	0.00	0.00	0.03	-0.02	0.08	0.03	-0.03	0.01
agef3	0.07	0.04	0.01	-0.02	-0.01	-0.03	-0.02	0.02	0.00	0.17
agef4	0.06	0.02	0.00	-0.01	-0.02	-0.06	0.00	0.02	0.03	-0.10
amot1	0.30	0.19	0.28	0.23	0.19	0.12	0.26	0.24	0.40	0.45
amot2	0.29	0.13	0.16	0.11	0.12	0.10	0.17	0.19	0.37	0.35
amot3	0.21	0.08	0.09	0.08	0.05	0.03	0.14	0.15	0.22	0.28
amot4	0.20	0.03	0.10	0.04	0.00	-0.01	0.07	0.08	0.08	0.41
int.stay	0.47	0.21	0.11	0.08	0.14	0.17	0.01	0.20	0.19	0.17
comorb	-0.03	0.00	0.06	0.02	0.02	0.01	0.13	0.04	0.01	0.25
readmit	-0.06	0.05	-0.03	0.05	0.07	0.02	0.16	0.00	0.11	0.57
<u>UDSmr LOS</u>										
constant	1.00	1.03	1.01	1.04	1.14	1.13	0.98	1.04	1.05	1.00
acog1	0.01	0.11	0.03	0.01	0.00	0.07	0.07	0.00	0.01	0.17
acog2	-0.03	0.11	0.03	0.02	0.01	0.05	0.04	0.00	0.08	0.11
acog3	-0.01	0.05	0.01	0.03	0.01	0.05	0.00	0.03	0.00	0.15
acog4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04
agef1	0.05	-0.03	-0.02	-0.01	0.00	-0.03	0.01	0.06	0.01	0.02
agef2	0.06	-0.02	-0.02	0.00	0.02	-0.03	0.08	0.02	-0.02	0.04
agef3	0.04	0.05	-0.01	-0.03	0.00	-0.02	0.05	0.01	0.05	0.15
agef4	0.07	0.02	-0.01	0.00	0.01	-0.01	0.06	0.02	0.02	0.05
amot1	0.28	0.18	0.29	0.26	0.21	0.07	0.38	0.25	0.46	0.51
amot2	0.32	0.12	0.18	0.13	0.12	0.02	0.26	0.20	0.40	0.30
amot3	0.17	0.06	0.10	0.10	0.05	-0.04	0.17	0.14	0.25	0.18
amot4	0.16	0.02	0.09	0.04	-0.01	-0.05	0.07	0.07	0.13	0.06
int.stay	0.29	0.21	0.18	0.11	0.19	0.20	0.18	0.19	0.18	0.06
comorb	-0.04	0.00	0.07	0.02	0.02	0.03	0.09	0.04	0.03	0.08
readmit	-0.06	0.06	0.00	0.04	0.05	0.02	0.20	0.01	0.12	0.17

Table C.2
T Statistics from Gold Standard Models

	RIC									
	01	02	03	04	05	06	07	08	09	10
<u>MEDPAR Cost</u>										
constant	548.9	103.4	136.6	71.2	159.8	199.6	432.1	462.6	205.6	195.8
acog1	2.6	3.5	2.3	-0.9	1.8	-0.5	5.4	14.3	1.8	2.2
acog2	5.3	2.5	2.0	0.3	2.5	0.6	8.2	7.2	2.9	1.4
acog3	7.0	2.0	1.8	-0.5	1.7	-0.9	6.2	0.0	2.5	1.2
acog4	5.0	1.5	1.9	0.0	0.0	0.7	0.0	0.0	0.0	0.0
agef1	10.7	4.0	1.7	0.1	1.6	3.0	0.3	-6.7	0.5	0.5
agef2	10.6	2.4	2.2	0.6	0.8	2.8	0.3	-5.6	0.0	1.4
agef3	7.9	1.4	1.5	0.6	1.0	1.7	2.6	-6.1	1.2	1.8
agef4	5.7	1.5	0.6	0.5	1.5	1.9	2.7	-2.5	1.1	-1.8
amot1	47.8	7.8	10.1	7.8	16.0	13.2	19.8	27.1	9.4	9.5
amot2	47.0	5.8	9.0	4.2	13.1	12.2	18.9	16.5	6.5	8.4
amot3	37.0	4.3	6.6	3.1	9.3	10.1	12.1	10.4	5.2	5.4
amot4	23.1	3.2	2.6	1.3	4.0	6.5	7.9	7.1	2.0	2.3
int.stay	12.8	4.9	5.7	2.2	5.9	6.3	9.0	8.6	5.5	9.6
comorb	-0.2	1.6	2.6	0.4	4.3	3.3	7.7	9.6	4.2	0.7
readmit	-2.3	-1.5	-2.3	-3.1	0.5	-0.4	2.2	0.1	1.8	-2.0
<u>MEDPAR LOS</u>										
constant	165.0	29.9	40.2	23.1	46.3	58.8	135.9	132.7	60.3	65.7
acog1	-1.5	3.9	2.6	-1.3	1.9	-0.4	3.4	13.7	1.5	0.6
acog2	1.7	2.5	1.6	0.4	1.7	0.5	7.4	8.6	3.0	1.6
acog3	4.1	2.4	1.5	-0.2	1.8	-0.5	6.8	0.0	3.5	0.6
acog4	3.3	2.0	1.7	0.0	0.0	0.9	0.0	0.0	0.0	0.0
agef1	9.1	3.6	-0.3	0.0	1.5	1.3	-3.8	-9.0	-0.9	-2.6
agef2	9.4	2.1	1.5	0.3	0.1	2.0	-3.0	-7.0	-0.8	-0.2
agef3	7.2	1.7	0.9	0.1	0.8	1.2	0.7	-7.3	0.3	-0.5
agef4	5.4	1.7	-0.6	0.3	1.2	1.8	1.2	-3.1	0.6	0.5
amot1	49.0	7.3	9.4	8.0	15.7	12.0	19.3	26.6	9.7	9.2
amot2	49.1	5.3	8.5	4.6	13.0	11.5	18.8	16.9	7.5	9.2
amot3	39.5	5.0	6.0	4.0	9.0	9.8	11.8	10.5	5.3	6.1
amot4	24.4	3.3	2.2	1.9	3.5	6.8	8.6	6.8	2.1	3.9
int.stay	13.8	5.0	4.9	2.9	6.3	6.4	10.0	8.9	5.0	8.7
comorb	-4.4	1.4	1.1	0.7	2.6	0.9	3.3	7.9	2.5	-0.4
readmit	-1.3	-1.7	-1.5	-3.0	1.5	-0.4	1.7	1.0	2.2	-2.3

Table C.2--continued

	RIC									
	01	02	03	04	05	06	07	08	09	10
UDSmr LOS	-									
constant	191.7	53.2	49.7	47.0	60.1	75.4	146.1	149.4	69.5	70.5
acog1	1.7	9.4	5.1	1.3	2.3	0.1	3.9	14.9	2.3	2.0
acog2	4.2	8.8	4.5	0.9	2.0	1.1	7.5	7.0	3.2	2.4
acog3	5.2	7.1	2.7	0.0	2.7	1.4	7.3	0.0	3.4	1.7
acog4	4.8	4.4	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
agef1	10.0	3.4	3.9	5.8	3.6	0.0	-5.2	-13.2	-1.9	-3.5
agef2	9.5	4.4	4.1	3.9	2.7	1.3	-3.1	-9.0	-1.3	-2.7
agef3	8.0	5.2	2.7	4.0	0.1	1.7	0.5	-6.7	0.3	-0.7
agef4	5.5	3.5	1.8	3.1	0.5	1.5	1.7	-4.7	1.6	0.6
amot1	63.0	25.3	17.5	24.3	23.0	16.0	23.6	30.2	12.0	11.9
amot2	61.4	20.0	13.9	17.3	18.0	14.1	21.2	17.7	8.7	10.7
amot3	50.0	12.9	9.6	14.3	12.0	11.7	13.9	11.8	6.4	6.3
amot4	29.2	6.8	4.8	7.5	5.6	7.0	8.7	7.5	3.4	5.0
int.stay	18.1	6.3	8.4	5.3	9.1	7.8	12.1	9.8	6.1	9.3
comorb	-4.9	-0.2	-0.4	-1.6	2.2	0.6	3.1	7.9	2.5	-0.1
readmit	-0.4	0.6	-2.1	-4.7	1.3	0.2	2.5	1.1	2.4	-1.9

	RIC									
	11	12	13	14	15	16	17	18	19	20
MEDPAR Cost										
constant	48.2	119.8	116.7	159.4	128.5	108.3	58.7	239.4	48.2	17.4
acog1	1.2	3.7	1.8	1.7	1.0	2.7	2.1	1.3	1.2	0.7
acog2	-1.2	4.5	2.3	1.0	0.8	3.3	1.6	1.1	1.0	1.3
acog3	-0.7	1.6	1.1	2.3	-0.5	0.9	1.9	1.9	0.0	0.3
acog4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6
agef1	1.1	0.9	1.2	0.1	2.4	-0.3	2.0	5.4	1.9	0.7
agef2	0.8	0.8	0.4	0.0	2.7	-0.6	1.4	3.7	0.0	0.1
agef3	1.0	2.2	0.1	-1.3	1.0	-0.6	-0.1	2.3	-0.2	0.9
agef4	1.7	1.9	-0.1	-0.7	0.2	-1.7	0.9	1.5	0.9	-0.2
amot1	3.4	6.3	8.9	13.5	6.8	3.4	5.0	16.7	6.9	2.6
amot2	4.1	3.9	5.9	8.4	3.9	2.6	2.8	13.7	4.7	2.3
amot3	2.5	2.4	3.5	6.1	1.9	0.6	2.8	9.8	3.1	1.8
amot4	2.6	0.4	3.5	2.5	-0.2	-0.6	0.9	5.0	1.1	2.3
int.stay	3.6	2.8	2.0	3.4	2.7	1.9	0.9	8.9	2.4	0.4
comorb	0.2	1.7	3.4	5.2	2.2	1.7	3.2	5.6	-0.3	1.4
readmit	-1.0	1.4	-1.3	1.3	2.4	-0.1	2.9	-1.1	0.7	1.5

Table C.2--continued

	RIC									
	11	12	13	14	15	16	17	18	19	20
<u>MEDPAR LOS</u>										
constant	13.8	33.5	32.1	52.0	44.8	37.3	17.8	71.0	14.2	4.2
acog1	0.4	3.7	1.5	1.4	0.2	2.4	1.0	-0.3	1.0	0.9
acog2	-1.0	4.8	1.8	1.6	0.0	3.4	0.6	0.0	1.1	1.8
acog3	-0.2	2.7	0.7	2.5	-0.4	2.1	1.0	1.0	0.0	1.2
acog4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6
agef1	-0.3	-1.0	0.6	0.3	1.0	-1.4	1.0	2.7	1.3	1.2
agef2	0.2	-0.7	0.1	0.1	1.3	-0.7	1.5	2.5	-0.5	0.1
agef3	0.9	1.4	0.3	-1.1	-0.2	-1.1	-0.3	1.4	0.0	1.0
agef4	0.8	0.9	0.0	-0.3	-0.7	-1.9	0.0	1.3	0.4	-0.6
amot1	3.9	6.5	9.7	11.9	7.5	4.1	4.9	17.1	5.0	3.1
amot2	4.1	4.9	5.2	6.2	5.0	3.3	3.2	14.1	4.4	2.6
amot3	2.9	2.9	3.1	4.5	2.3	1.0	2.6	10.5	2.6	2.0
amot4	2.9	1.2	3.4	2.5	-0.2	-0.3	1.3	5.7	0.9	3.3
int.stay	3.9	3.7	1.6	2.6	3.9	2.2	0.1	8.5	2.1	0.6
comorb	-0.7	0.0	2.3	1.8	1.1	0.3	2.9	4.0	0.1	2.0
readmit	-0.8	1.7	-0.9	1.9	2.2	0.8	2.1	0.0	1.2	2.1
<u>UDSmr LOS</u>										
constant	14.9	36.5	34.7	54.8	46.1	39.1	27.8	77.2	19.6	22.6
acog1	0.2	4.4	1.4	0.8	0.1	3.2	2.2	-0.1	0.2	4.0
acog2	-0.7	5.4	1.4	1.7	0.3	2.4	1.4	-0.4	2.1	2.7
acog3	-0.3	2.0	0.6	2.3	0.6	2.1	0.0	2.4	0.0	3.9
acog4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.1
agef1	0.9	-1.2	-0.8	-0.3	0.1	-1.1	0.3	4.9	0.3	0.5
agef2	0.9	-0.9	-0.5	-0.1	1.1	-1.3	2.3	1.7	-0.4	0.9
agef3	0.7	1.9	-0.2	-1.7	0.2	-0.7	1.4	1.1	1.0	3.6
agef4	1.2	0.7	-0.3	-0.2	0.3	-0.4	1.9	1.8	0.4	1.3
amot1	4.5	6.5	10.7	14.0	8.5	2.7	10.8	19.8	8.7	11.7
amot2	5.5	4.5	6.6	7.5	5.4	0.8	8.0	15.8	7.8	7.5
amot3	2.6	2.5	3.9	5.7	2.3	-1.6	4.9	11.2	5.0	4.7
amot4	2.9	0.6	3.3	2.5	-0.6	-2.1	2.2	5.5	2.4	1.6
int.stay	3.5	4.1	3.2	3.9	5.4	3.0	3.2	9.6	2.7	1.0
comorb	-0.8	0.0	2.5	1.6	1.1	0.7	2.0	4.0	0.5	0.9
readmit	-1.0	2.2	0.0	1.4	1.7	0.7	3.7	0.5	1.6	2.1

Appendix D

MEMBERS OF TECHNICAL ADVISORY PANELS

I. Clinical and Functional Status Panel

Bruce H. Dobkin, MD
Professor of Neurology, UCLA School of Medicine
Medical Director, UCLA Neurologic Rehabilitation and Research Unit
Medical Director, UCLA Comprehensive Rehabilitation Program

Pamela W. Duncan, PhD, PT
Associate Professor, Health Services Administration, University of Kansas
Director of Research, Center on Aging, University of Kansas Medical Center
Program Director, Health Services Research, Kansas City and Leavenworth VAs

Byron B. Hamilton, MD, PhD
Director, Rehabilitation Services Research and Development Unit, Durham VAMC
Associate Clinical Professor, Duke University Medical Center

Richard Harvey, MD
Chief Executive Officer, Rehabilitation Medicine Clinic, South Carolina
Executive Vice President, Medical Services, Marianjoy Rehabilitation Hospital and Clinics
President, Rehabilitation Foundation Inc.
Professor, Physical Medicine and Rehabilitation, Rush Medical College

Vincent Mor, PhD
Professor of Medical Science, Brown University
Director, Center for Gerontology and Health Care Research, Brown University

Ex-Officio Members

William Buczko, PhD
Project Officer,
Office of Research and Demonstrations
Health Care Financing Administration

Patricia Brooks
Branch Chief, Coding
Health Care Financing Administration

II. Classification and Payment System Design Panel

Stuart Guterman
Deputy Executive Director
Prospective Payment Assessment Commission

Mark Hornbrook, PhD
Program Director, Health Services, Social, and Economic Studies
Kaiser Permanente Center for Health Research, Portland Oregon

John Melvin, MD
Professor and Deputy Chairman, Department of Physical Medicine &
Rehabilitation
Temple University
Vice President, Medical Affairs
Moss Rehabilitation Hospital

Elliot Roth, MD
Professor and Chairman, Physical Medicine and Rehabilitation
Northwestern University Medical School
Medical Director, Rehabilitation Institute of Chicago

Donald Schneider, PhD
Associate Professor, School of Management, Rensselaer Polytechnic
Institute
Director, Health System and Management Program, Rensselaer Polytechnic
Institute

Ex-Officio Members

William Buczko, PhD
Project Officer,
Office of Research and Demonstrations
Health Care Financing Administration

Patricia Brooks
Branch Chief, Coding
Health Care Financing Administration

REFERENCES

- Breiman, L. J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, Wadsworth, Inc., Belmont, CA, 1984.
- Carter, G. M., J. D. Rumpel, *The Interaction Between Payment Adjustors and the Size of the Outlier Pool under Medicare's Prospective Payment System*, RAND, MR-192-HCFA, 1993.
- Carter, G. M., Joan L. Buchanan, Tenzing Donyo, Moira Inkelas, and Karen L. Spritzer, *A Prospective Payment System for Inpatient Rehabilitation*, RAND, MR-867, 1997.
- Fiedler, R. C., C. V. Granger, and K. J. Ottenbacher, "The Uniform Data System for Medical Rehabilitation", *Am. J. of PM&R*, 75(2): 125-129, 1996.
- Freeman, J. L., R. B. Fetter, H. Park et al., "Diagnosis-Related Group Refinement with Diagnosis and Procedure-Specific Comorbidities and Complications," *Medical Care*, 33(8):806-827, 1995.
- Hamilton, B. B., C. V. Granger, F. S. Sherwin, M. Zielesny, and J. S. Tashman, "A Uniform National Data System for Medical Rehabilitation," Chapter 10, *Rehabilitation Outcomes: Analysis and Measurement* (M. J. Fuhrer, ed.), Paul H. Brookes Publishing Co., Baltimore, 1987.
- Hosek, S. D., R. Kane, M. Carney et al., *Charges and Outcomes for Rehabilitative Care: Implications for the Prospective Payment System*, RAND, R-3424-HCFA, 1986.
- Jaro, M. A., "Probabilistic Linkage of Large Public Health Data Files," *Statistics in Medicine*, 14:491-498, 1995.
- Kramer, A. M., T. B. Eilertsen, C. A. Hrinkevich, and R. E. Schlenker, *Rehabilitation of Medicare Patients in Rehabilitation Hospitals and Skilled Nursing Facilities*, Study Paper 5, Center for Health Services Research, University of Colorado, Denver, 1994.
- Newhouse, Joseph P., Shan Cretin, and Christina J. Witsberger, "Predicting Hospital Accounting Costs," *Health Care Financing Review*, Vol. 11, No. 1, Fall 1989, pp. 25-33.
- ProPAC, *Medicare and the American Health Care System*, Report to the Congress, Washington, D.C., June 1993.
- ProPAC, *Report and Recommendations to the Congress*, Washington, D.C., March 1996.



Stineman, M. G., J. A. Shea, A. Jette et al., "The Functional Independence Measure: Tests of Scaling Assumptions, Structure, and Reliability Across 20 Diverse Impairment Categories," *Archives of Physical Medicine and Rehabilitation*, 77:1101-1108, 1996.

Stineman, M. G., J. J. Escarce, J. E. Goin et al., "A Case Mix Classification for Medical Rehabilitation," *Medical Care*, 32(4):366-379, 1994.